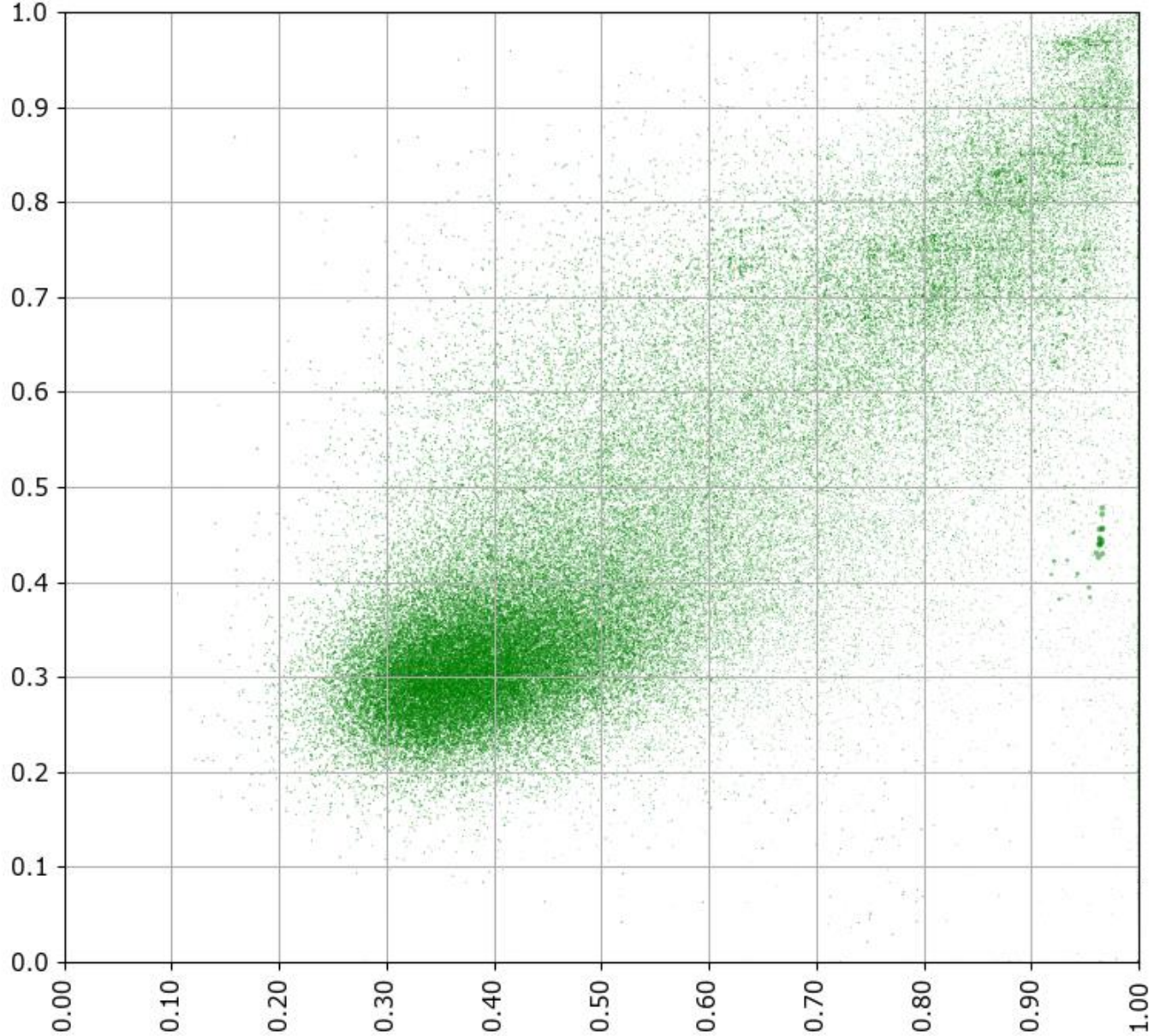

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

Что мы знаем о выборах

Федеральные выборы:

- Итоги
  - 10-1000 чисел
- Физическое наблюдение
  - тысячи - десятки тысяч наблюдателей
- Результаты по участкам
  - 1 - 10 миллионов чисел
- Видеонаблюдение*
  - ~1 млн часов

*2012 - †2021

Методы контроля

- Физическое наблюдение
  - Высокая трудоемкость, простая интерпретация, локальные данные
- Статистика
  - Невысокая трудоемкость, сложная интерпретация, глобальные данные
- Видеонаблюдение
  - Высокая трудоемкость, простая интерпретация, глобальные данные

Методы контроля

- Физическое наблюдение
  - Показывает наличие нарушений
- Статистика
  - Оценивает масштаб нарушений
- Видеонаблюдение
  - Подтверждает результаты физического наблюдения и статистики

Иерархия избирательной системы

- ЦИК РФ
- 85 субъектов федерации или 225 избирательных округов
- ~2800 территориальных избирательных комиссий
- Почти 100 тысяч избирательных участков
- ~110 миллионов избирателей
- Типичный избирательный участок:
  - ~1500-2500 избирателей (городской)
  - ~500 избирателей (сельский)
  - 10-300 избирателей (специальный)

- Все данные – на сайте izbirkom.ru

Иерархия избирательной системы

- ЦИК РФ
- 85 субъектов федерации или 225 избирательных округов
- ~2800 территориальных избирательных комиссий
- Почти 100 тысяч избирательных участков
- ~110 миллионов избирателей
- Типичный избирательный участок:
  - ~1500-2500 избирателей (городской)
  - ~500 избирателей (сельский)
  - 10-300 избирателей (специальный)

- Все данные – на сайте izbirkom.ru

Данные

Основные:

- Данные протоколов УИК
- Динамика явки в контрольных точках
- Данные о прикреплении/откреплении избирателей

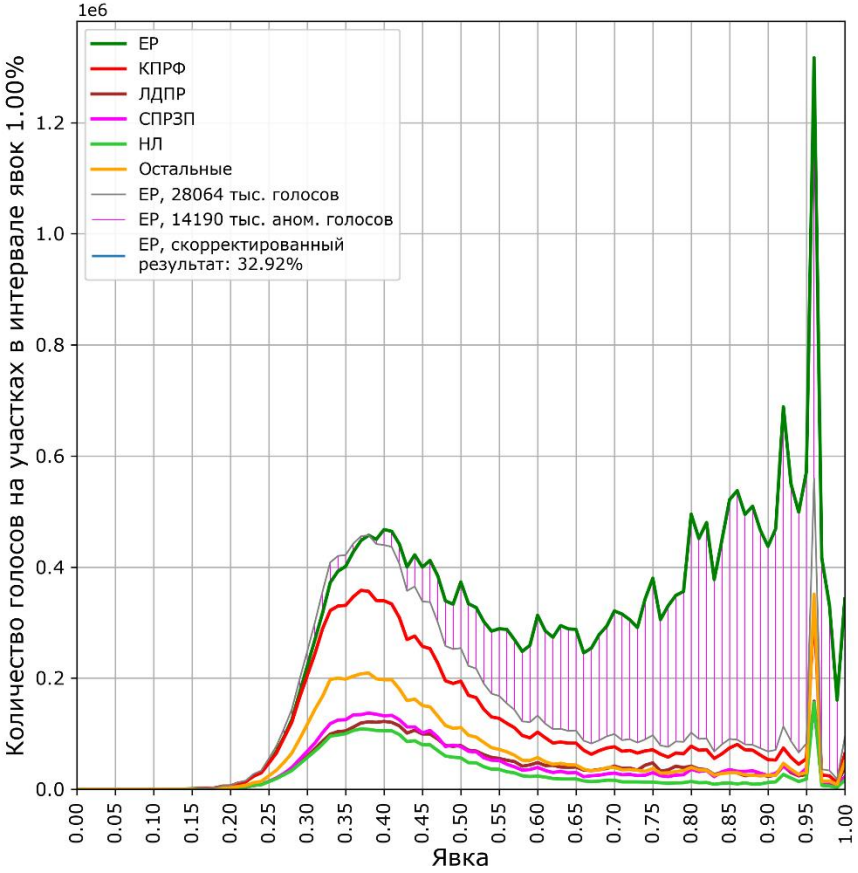
Дополнительные:

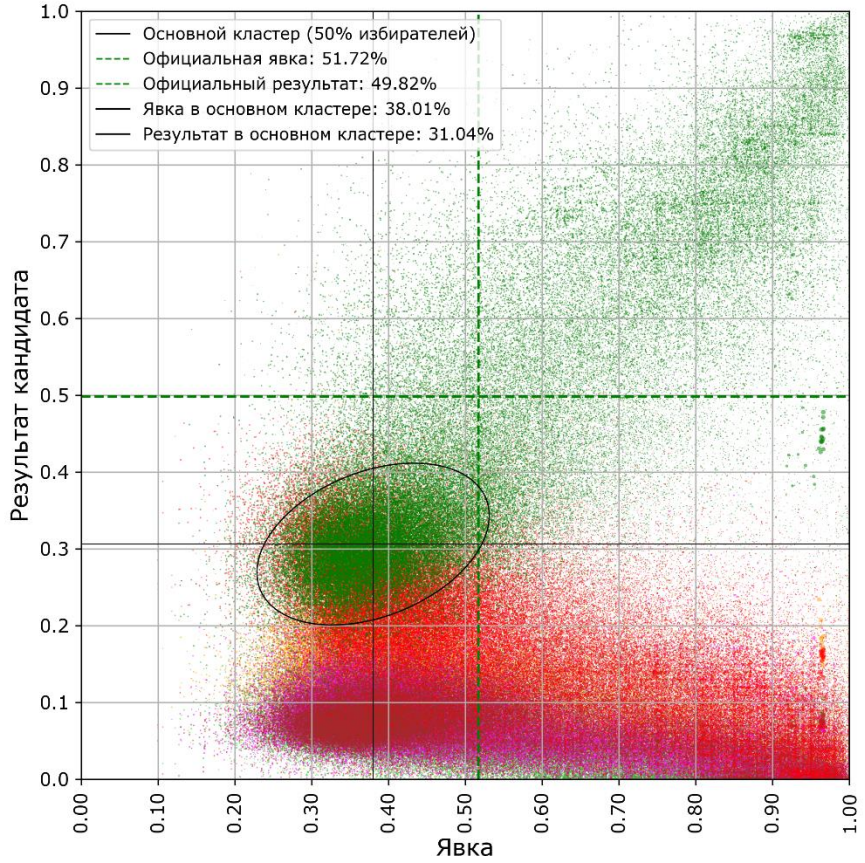
- История голосований и явки
- Геоданные участков (местоположение, совместное размещение)
- Привязка жилых домов к участкам

Выборы как социологический эксперимент

- 110 миллионов участников
- Каждый участок - локальный срез общества (типичный размер выборки при опросе – 1600 человек, городской избирательный участок – 1500–2500 человек).
- Результаты голосования на отдельно взятом избирательном участке – сумма случайных решений отдельно взятых избирателей
- Каждый избирательный участок вписан в многомерное пространство (географическое расположение, локальное расположение, историческое время, локальное время, параллельные выборы)
- поэтому «наивные» манипуляции с данными трудно скрыть

Вся Россия
Выборы в Государственную думу 2021 г.

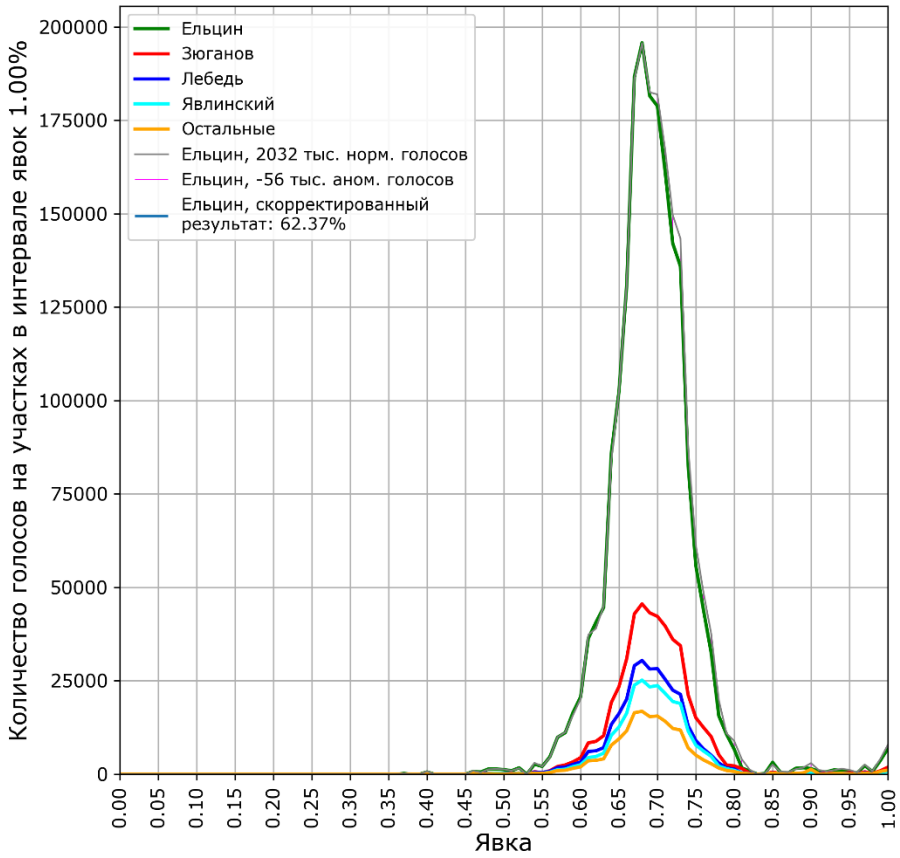


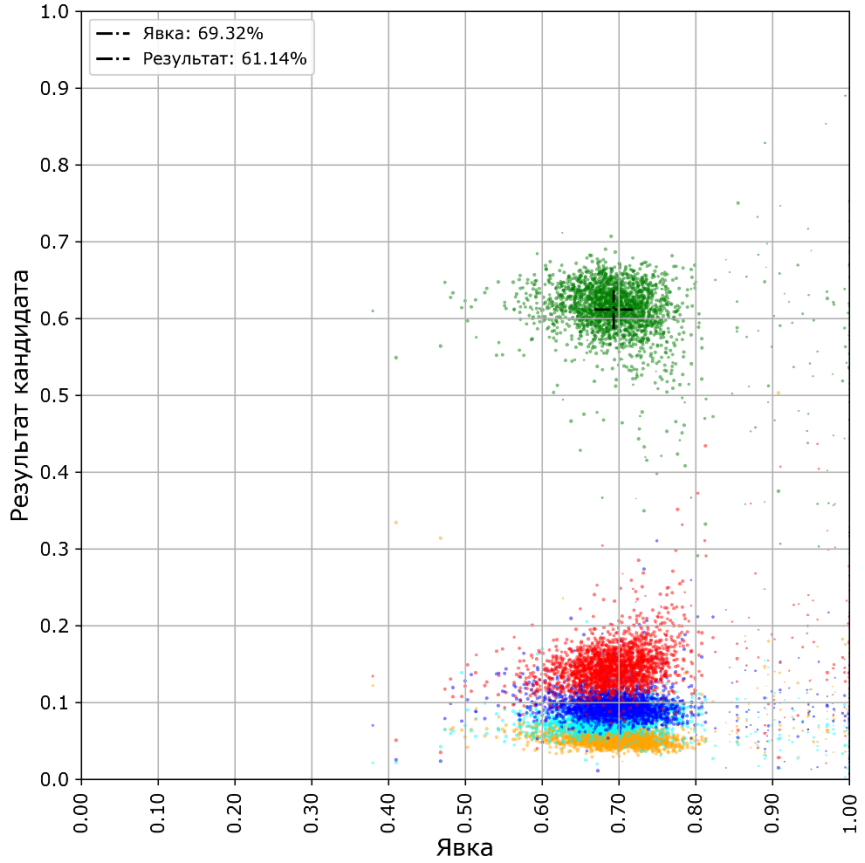


Загадочный фактор

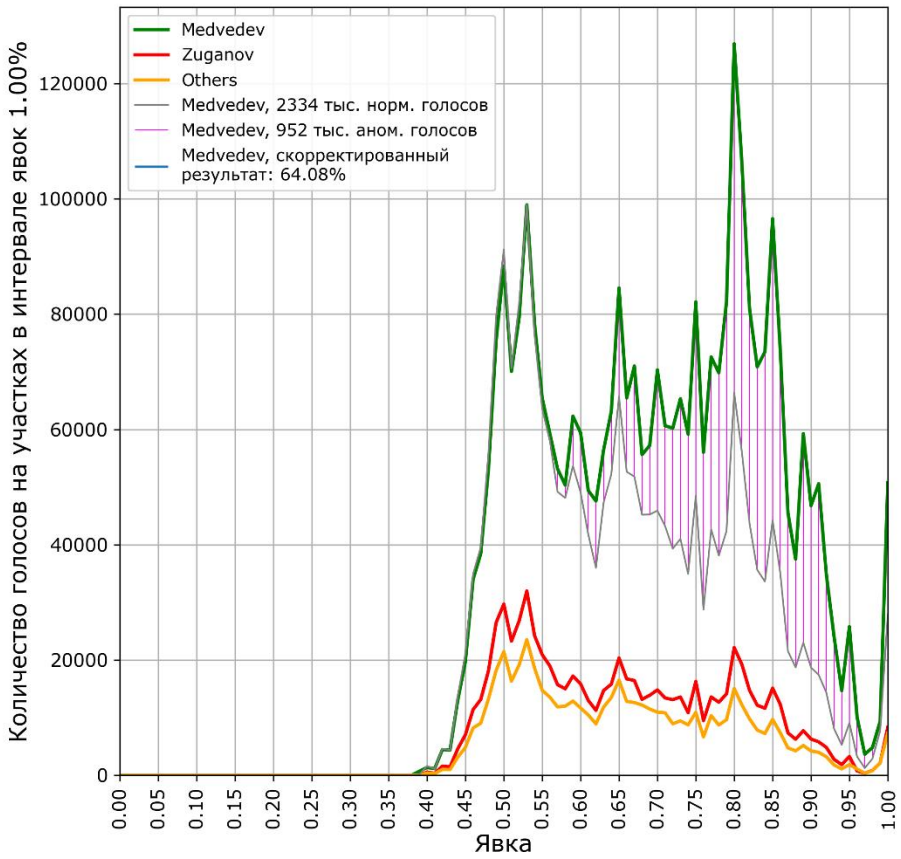
- Превосходит по силе все остальные факторы вместе взятые
- Действует на избирательные участки выборочно
- Повышает явку и результат официального кандидата
- Резко меняется со временем

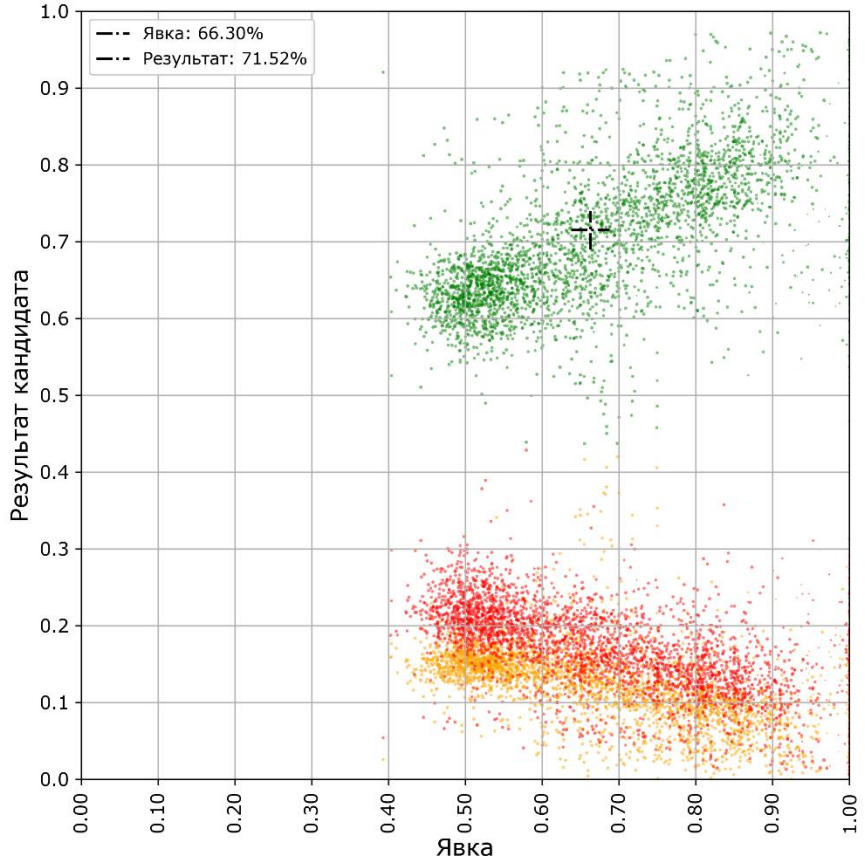
город Москва
Russia President election 1996_1



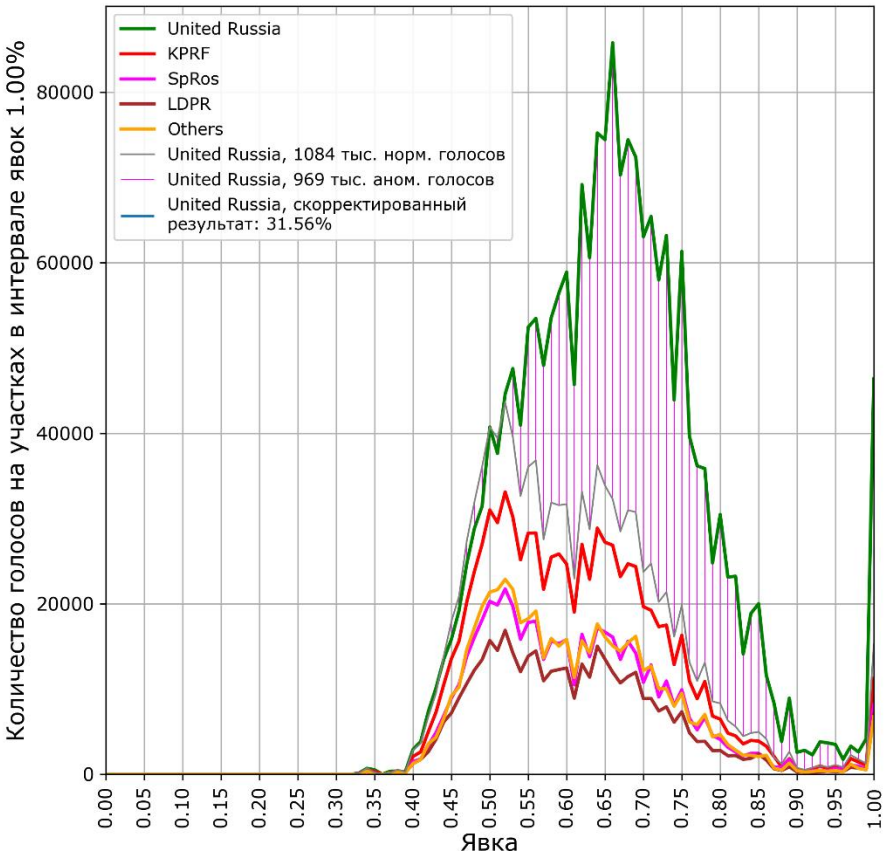


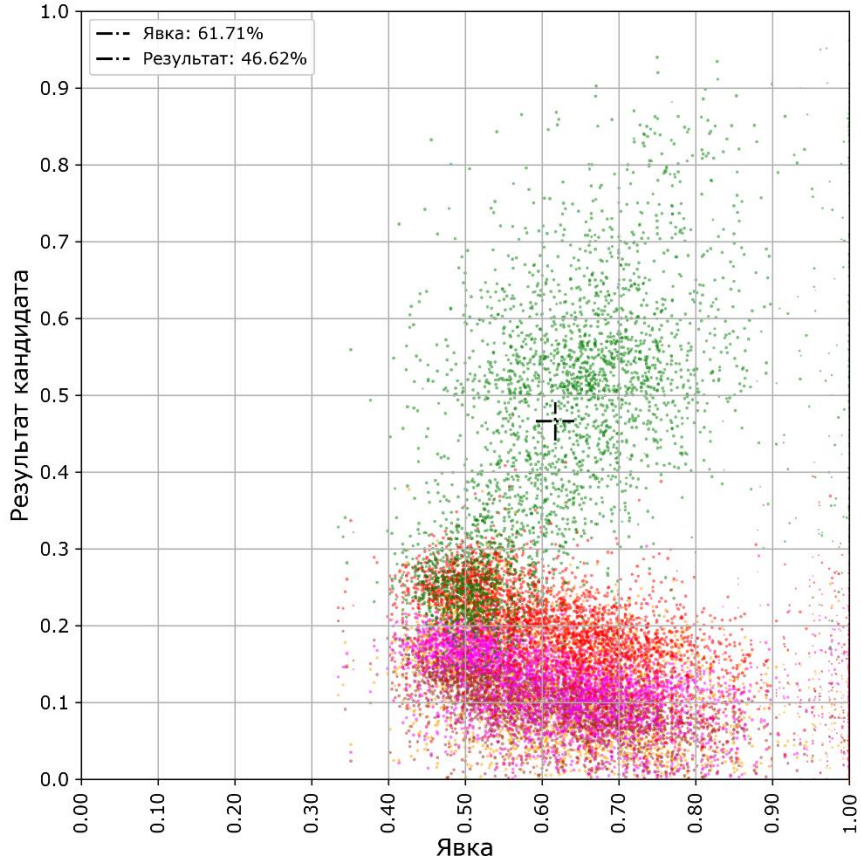
Город Москва
Russia President election 2008



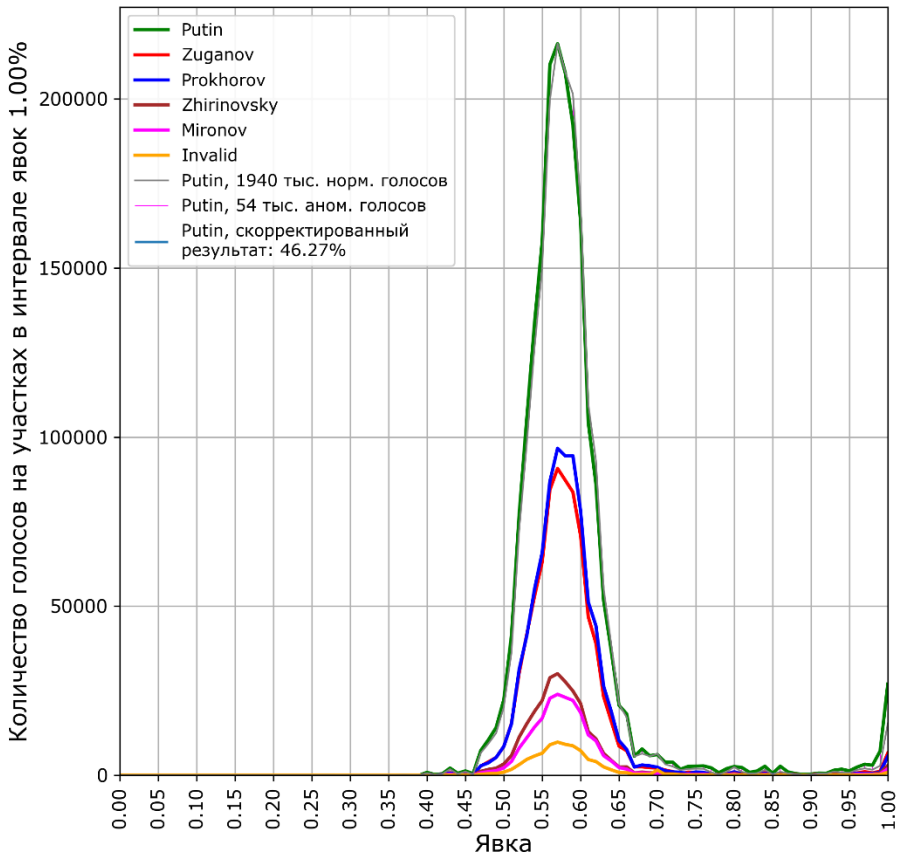


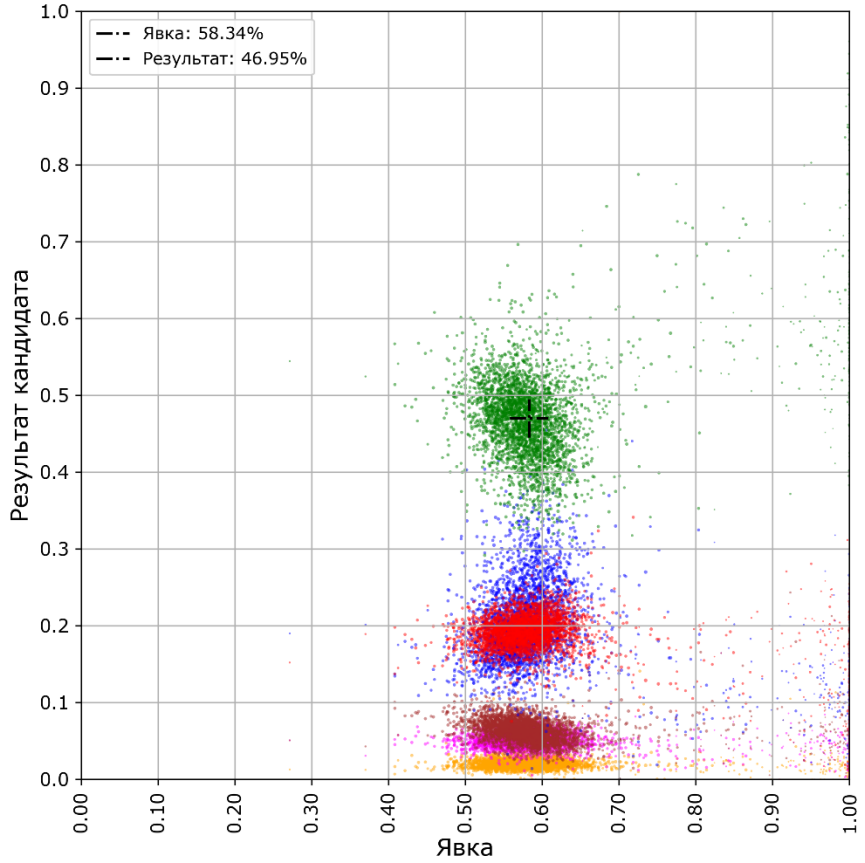
Город Москва
Russia Duma election 2011



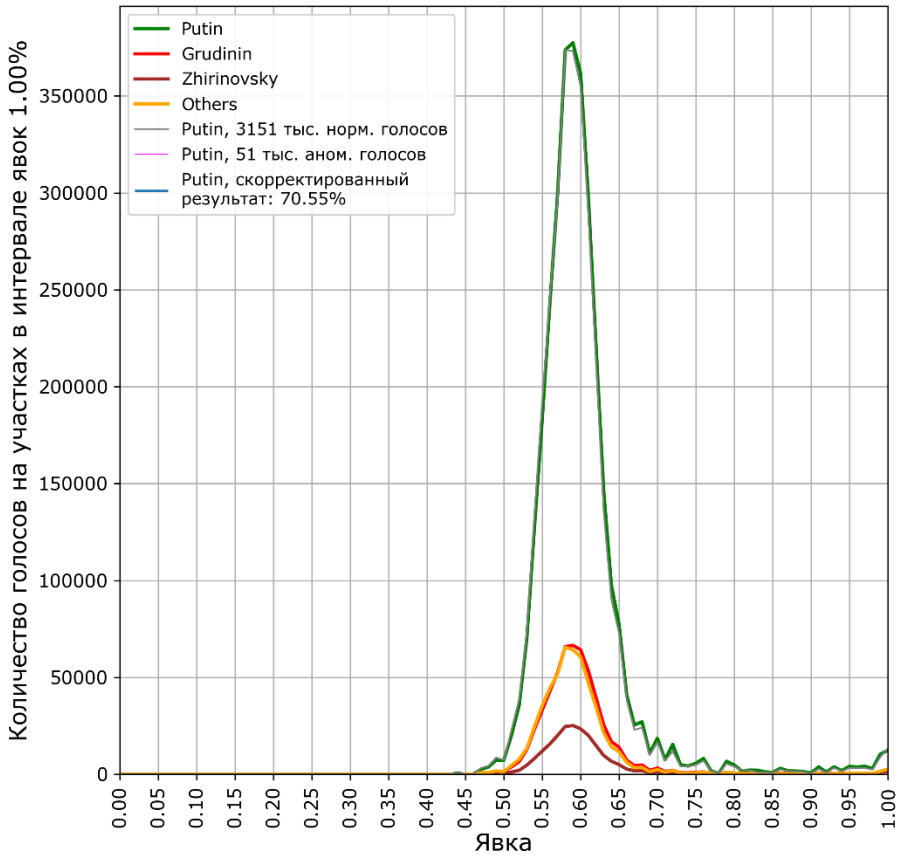


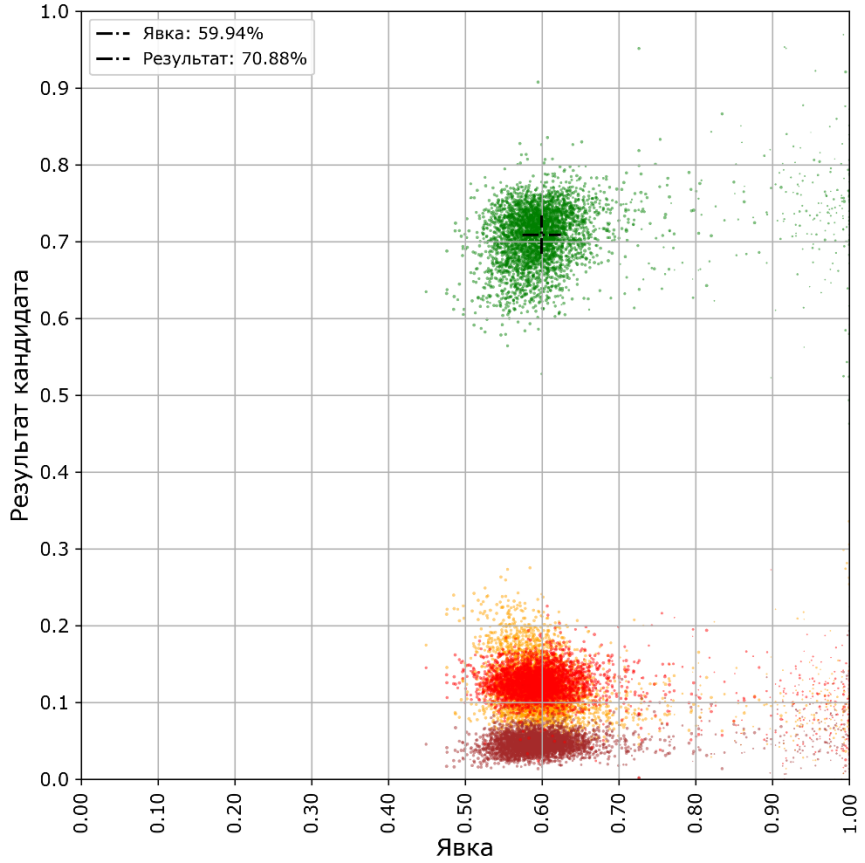
Город Москва
Russia President election 2012



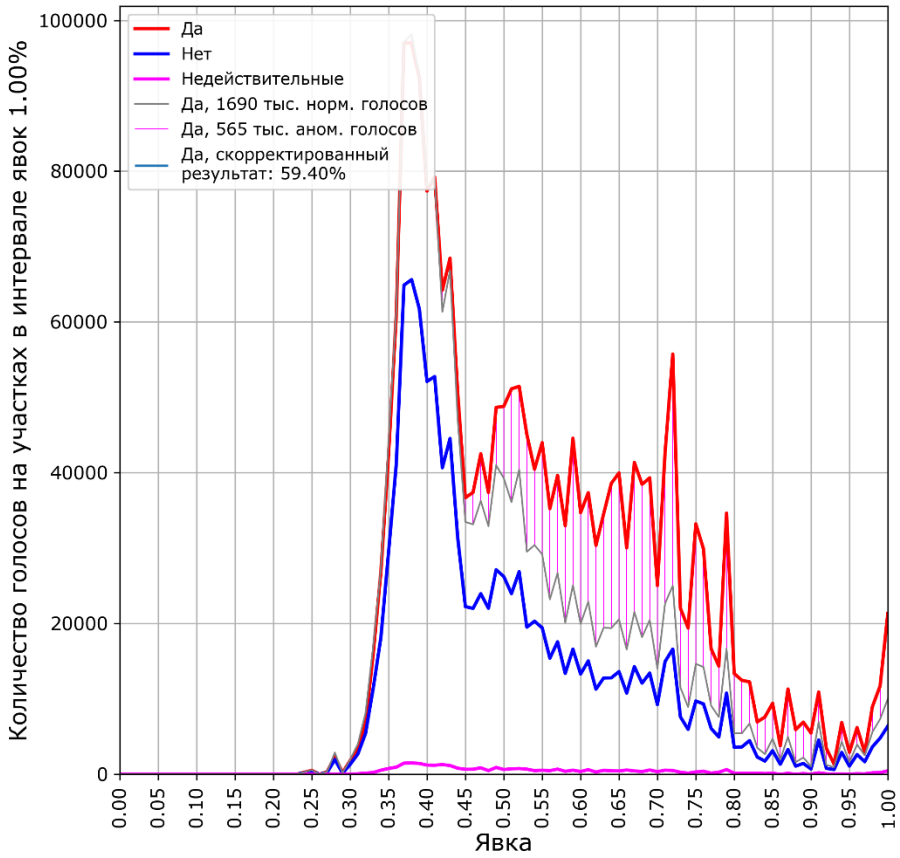


город Москва
Russia President election 2018

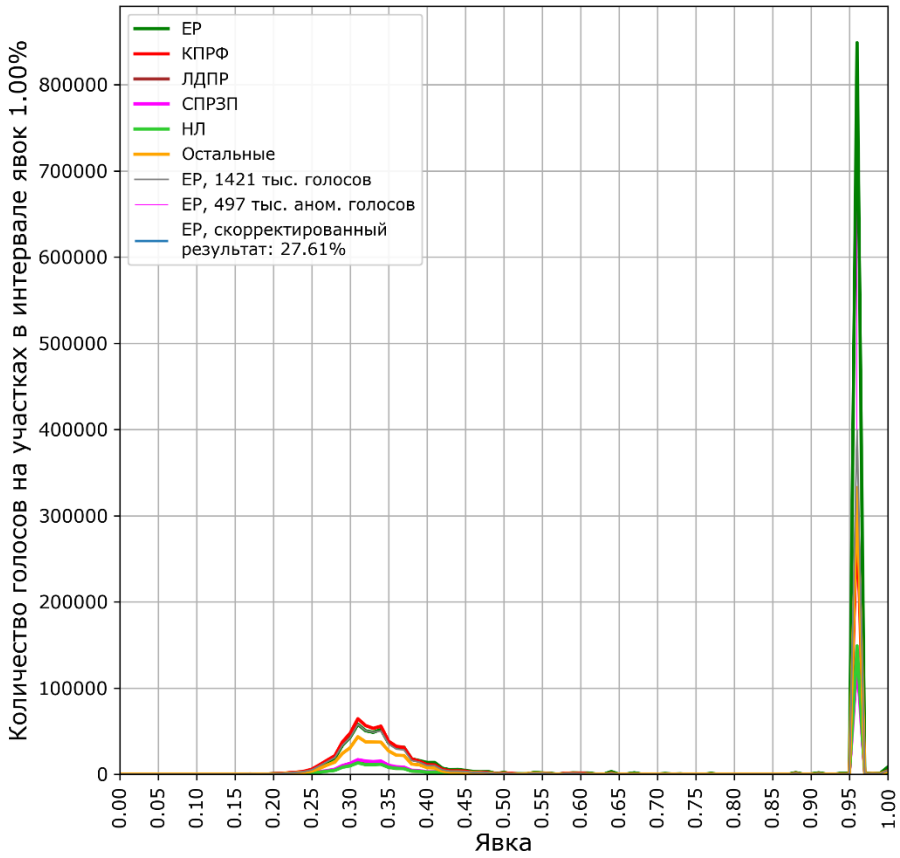


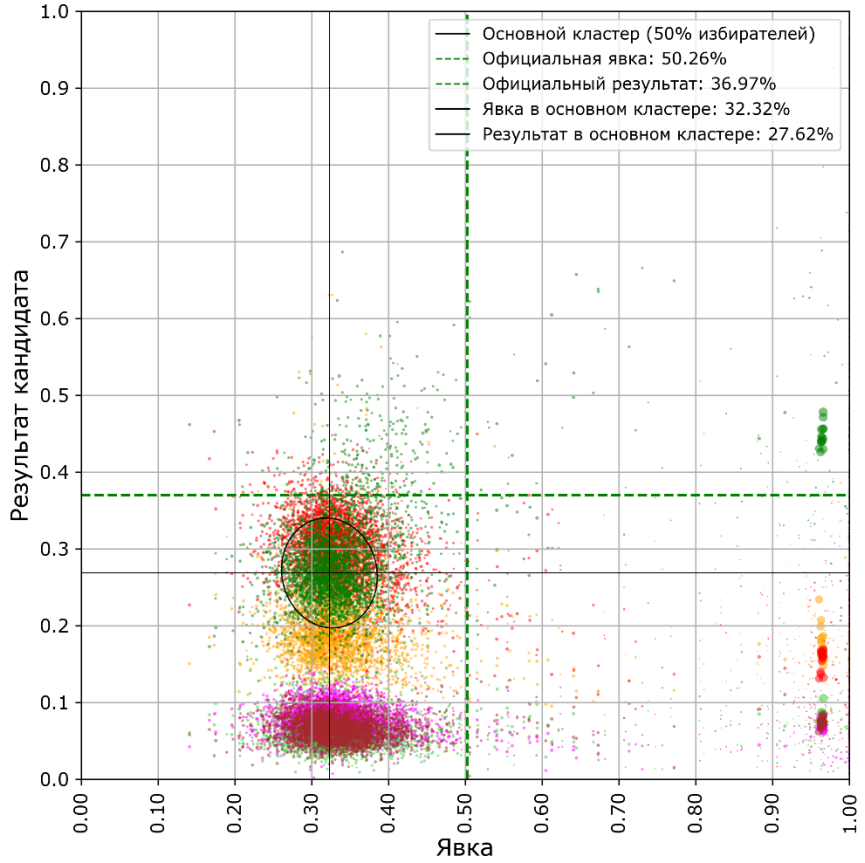


город Москва
Референдум 2020



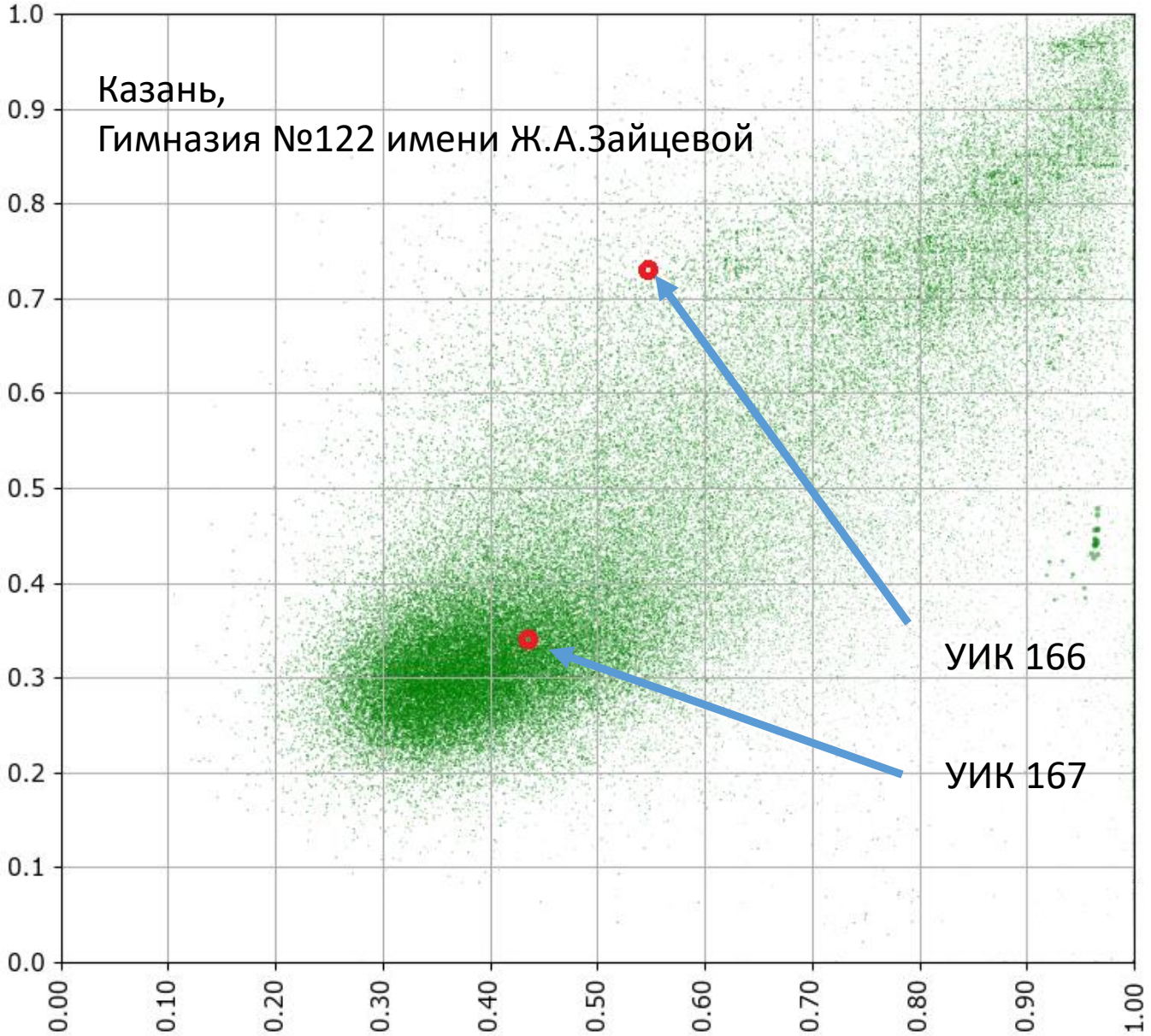
город Москва
Выборы в Государственную думу 2021 г.

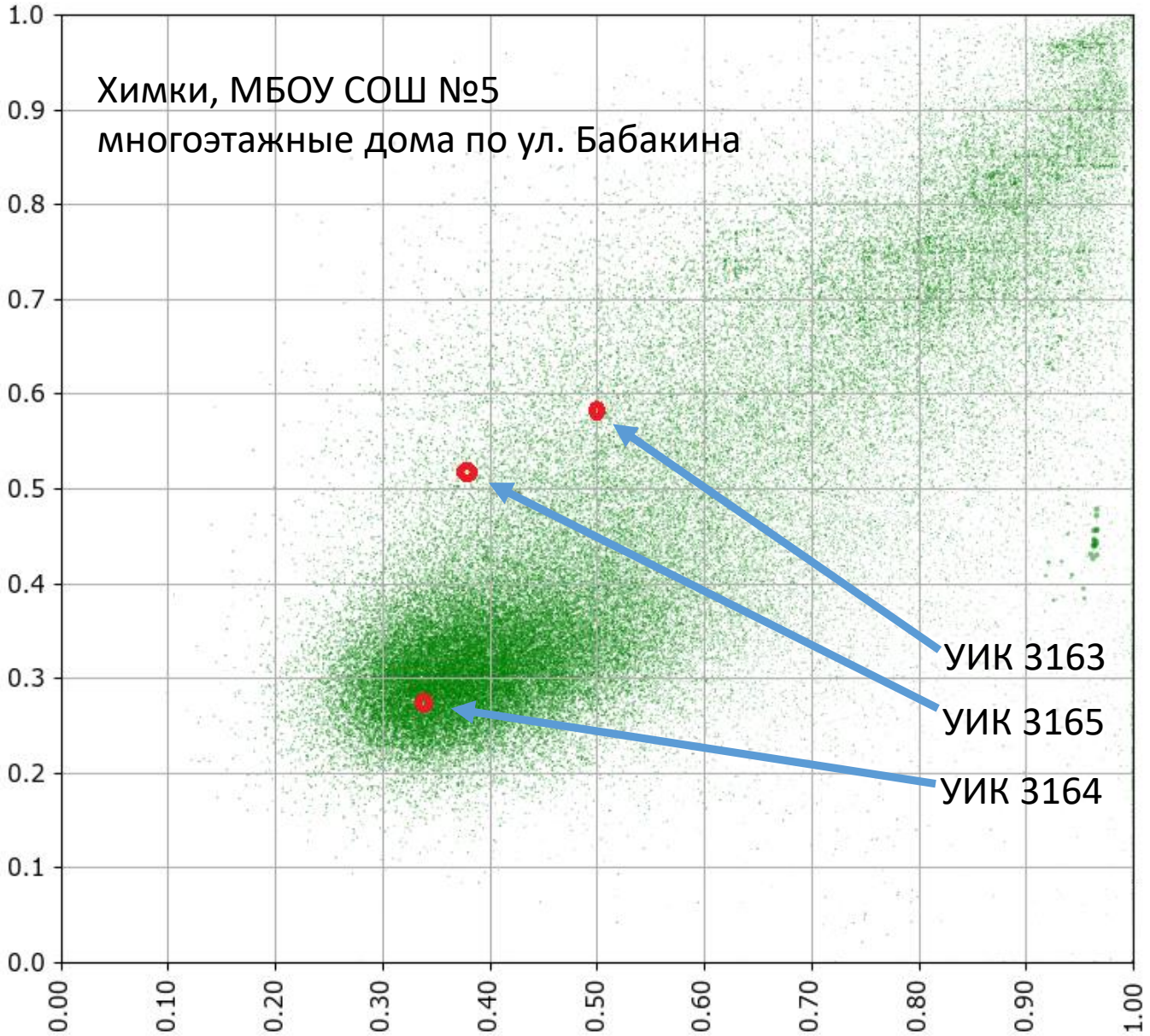




Загадочный фактор

- Резко меняется в пространстве

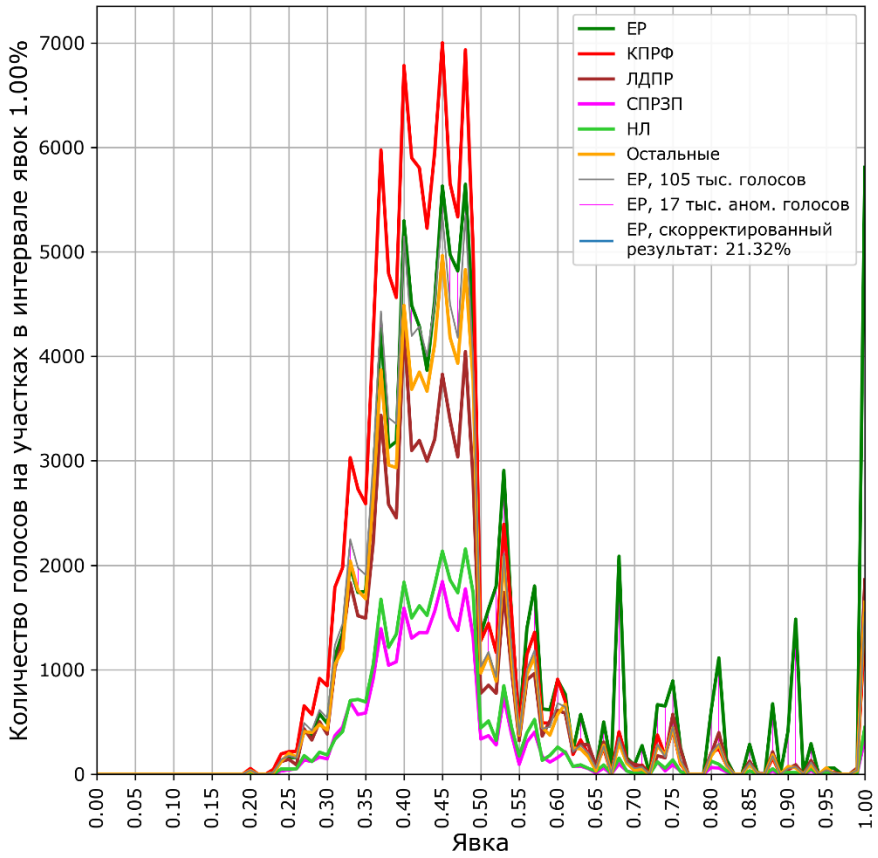

Казань,
Гимназия №122 имени Ж.А.Зайцевой
УИК 166
УИК 167
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

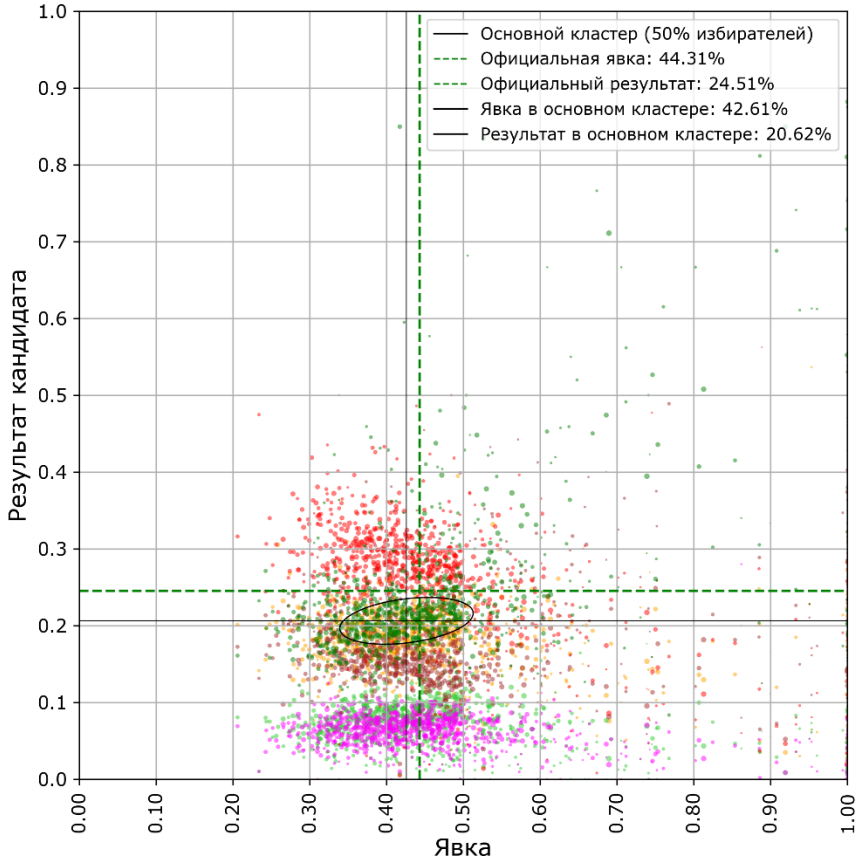

Химки, МБОУ СОШ №5
многоэтажные дома по ул. Бабакина
УИК 3163
УИК 3165
УИК 3164
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

Загадочный фактор

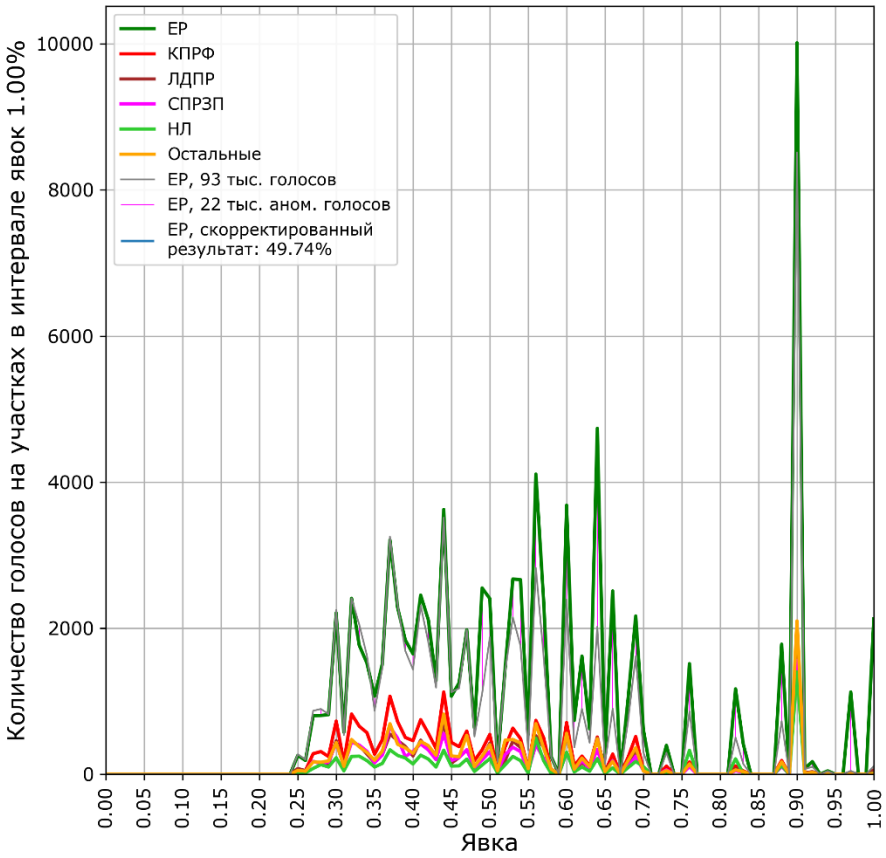
- Не связан с лояльностью или сельскостью

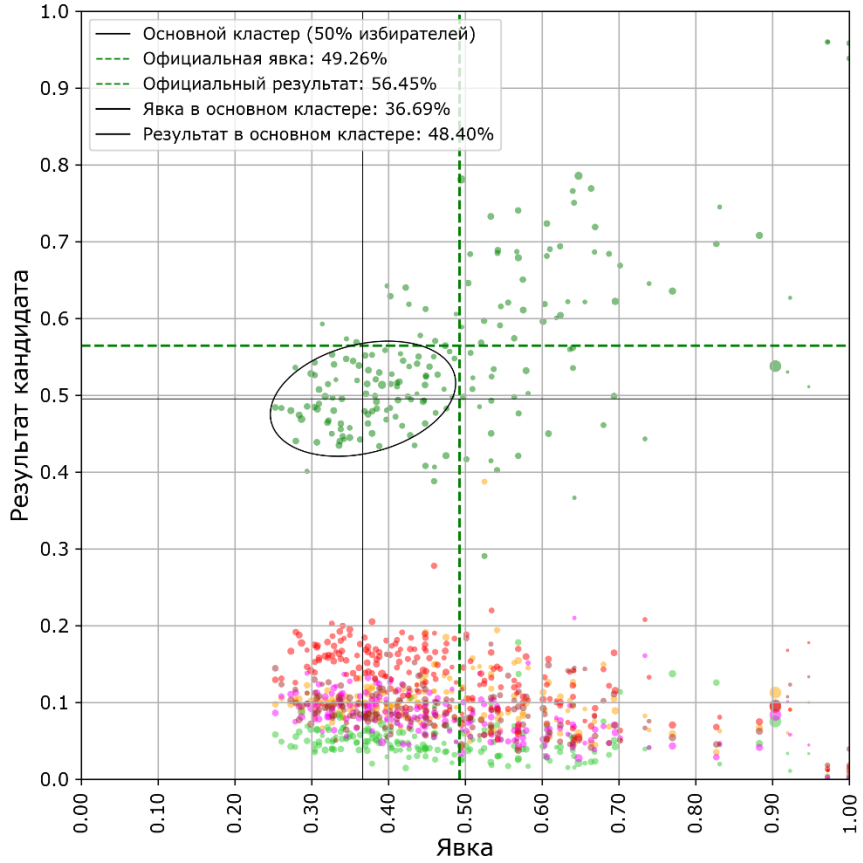
Хабаровский край
Выборы в Государственную думу 2021 г.



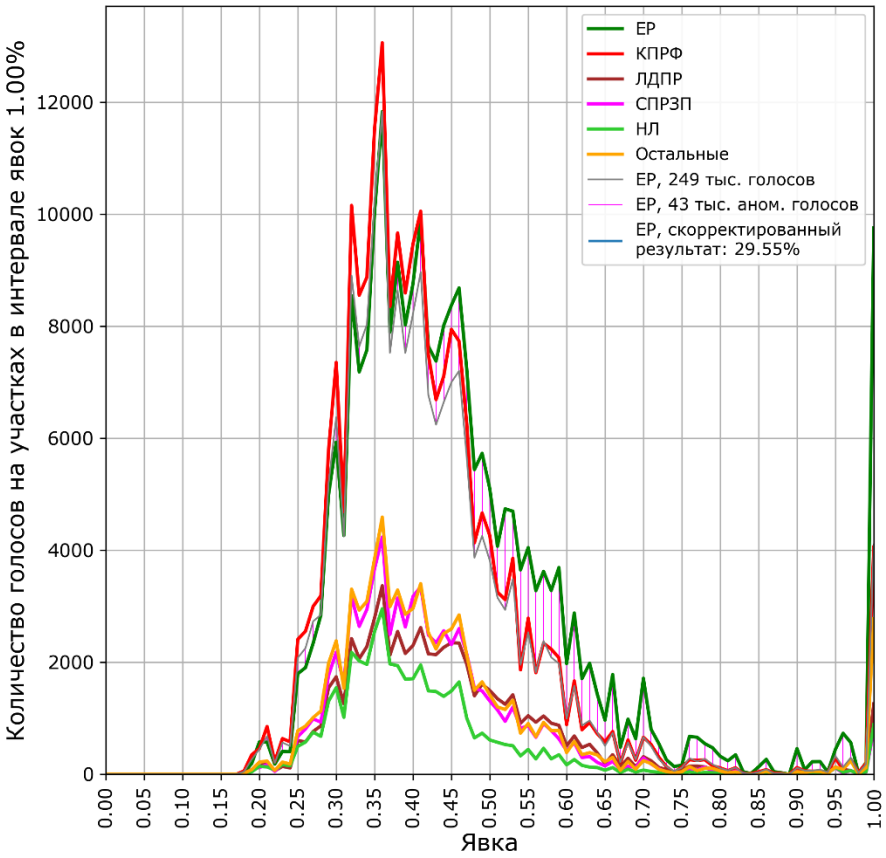


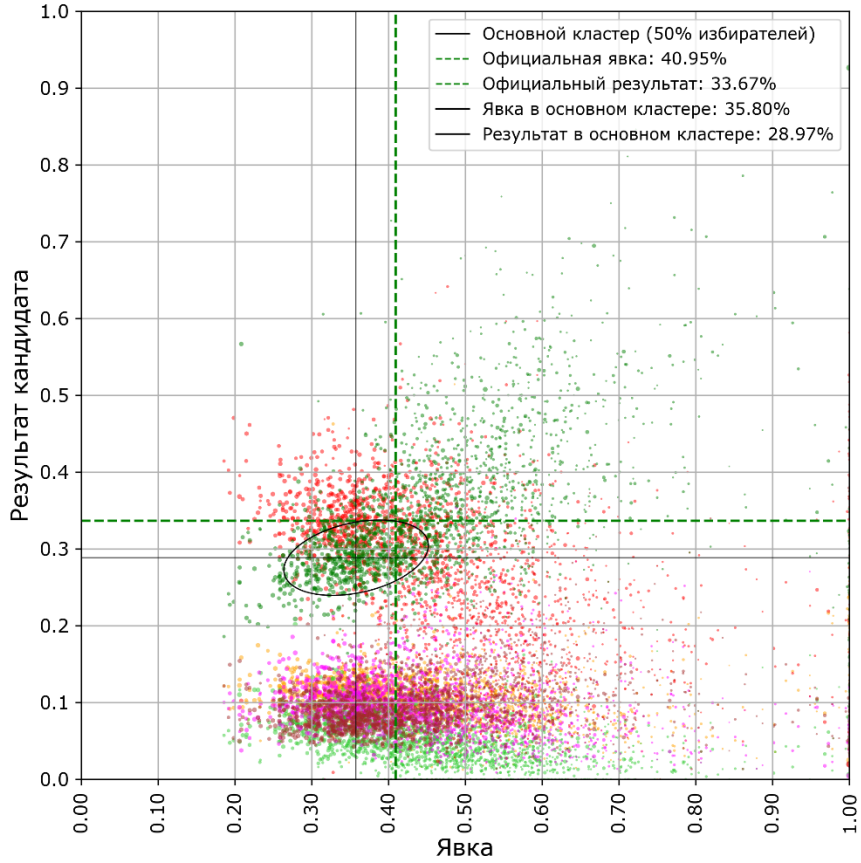
город Севастополь
Выборы в Государственную думу 2021 г.





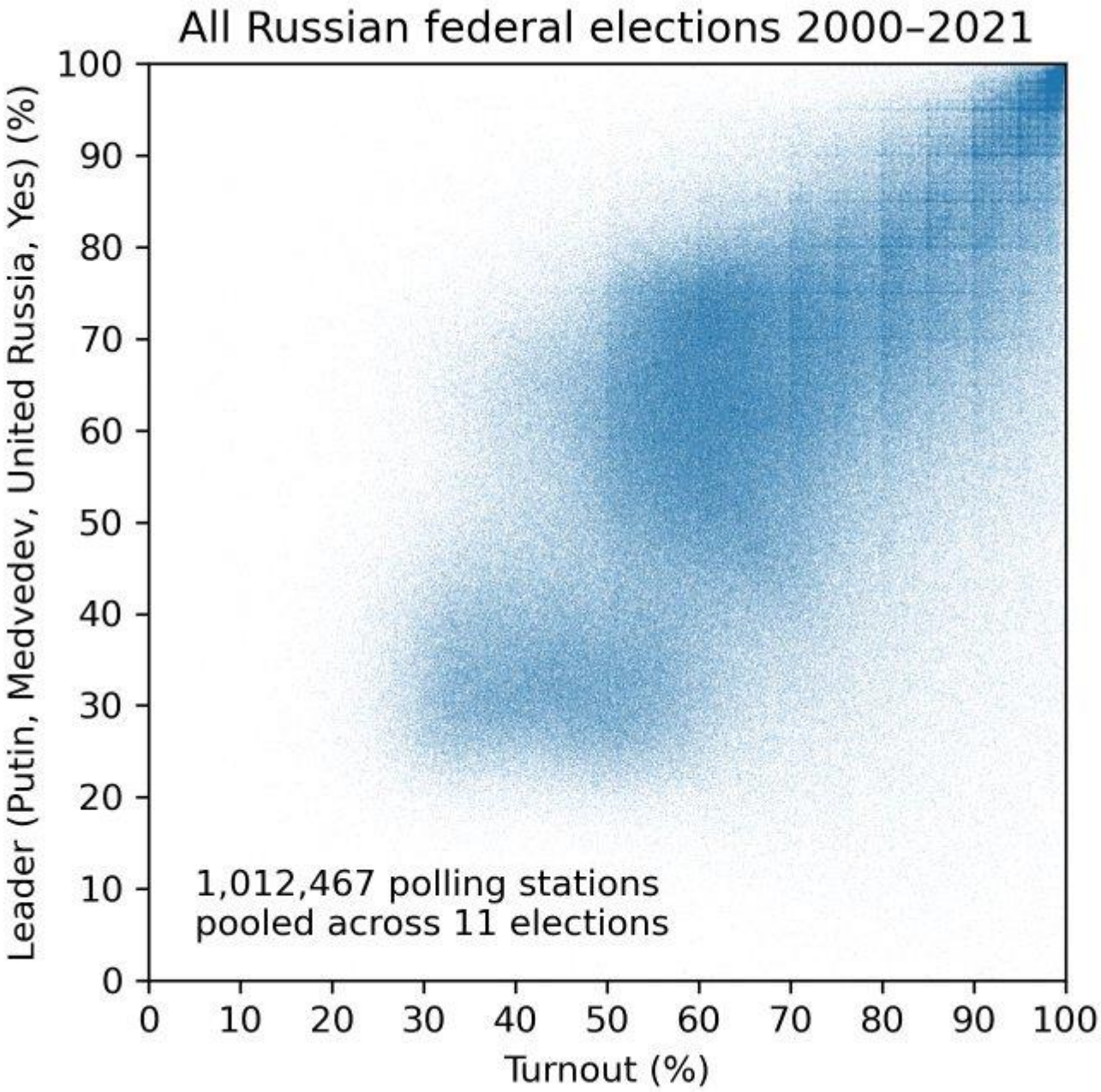
Алтайский край
Выборы в Государственную думу 2021 г.





Загадочный фактор

- Порождает загадочные структуры



(Д. Кобак https://twitter.com/hippopedoid, 2021)

WTF?

- (What is the factor)?

WTF?

- Административный ресурс?

- Явка – отчетный параметр (KPI), поэтому она портится первой
- Явка – хороший индикатор и сепаратор фальсификаций

Реальные результаты и фальсификации – точки формирования

Реальные результаты голосования складываются из действий отдельных избирателей (сотни и тысячи человек для каждого избирательного участка)

- имеют естественный статистический разброс и понятное распределение
- зависят от обнаружимых и объяснимых факторов

Реальные результаты и фальсификации – точки формирования

Фальсифицированные результаты образуются волей небольшого числа лиц (единицы для каждого избирательного участка)

- часто имеют неестественные статистические характеристики
- связаны не с реальностью, а с пожеланиями начальства (или с представлениями действующих лиц о них)

Модель избирательного участка

- 2000 избирателей

- Каждый избиратель бросает монетку:
  - С вероятностью 0.6 идет голосовать
  - С вероятностью 0.4 остается дома
- Можно описать случайной переменной X:
  - С вероятностью 0.6: X=1
  - С вероятностью 0.4: X=0

- Итоговое количество проголосовавших – сумма независимых случайных переменных Xi:

$$T = \Sigma X_i,\ i=1....2000$$

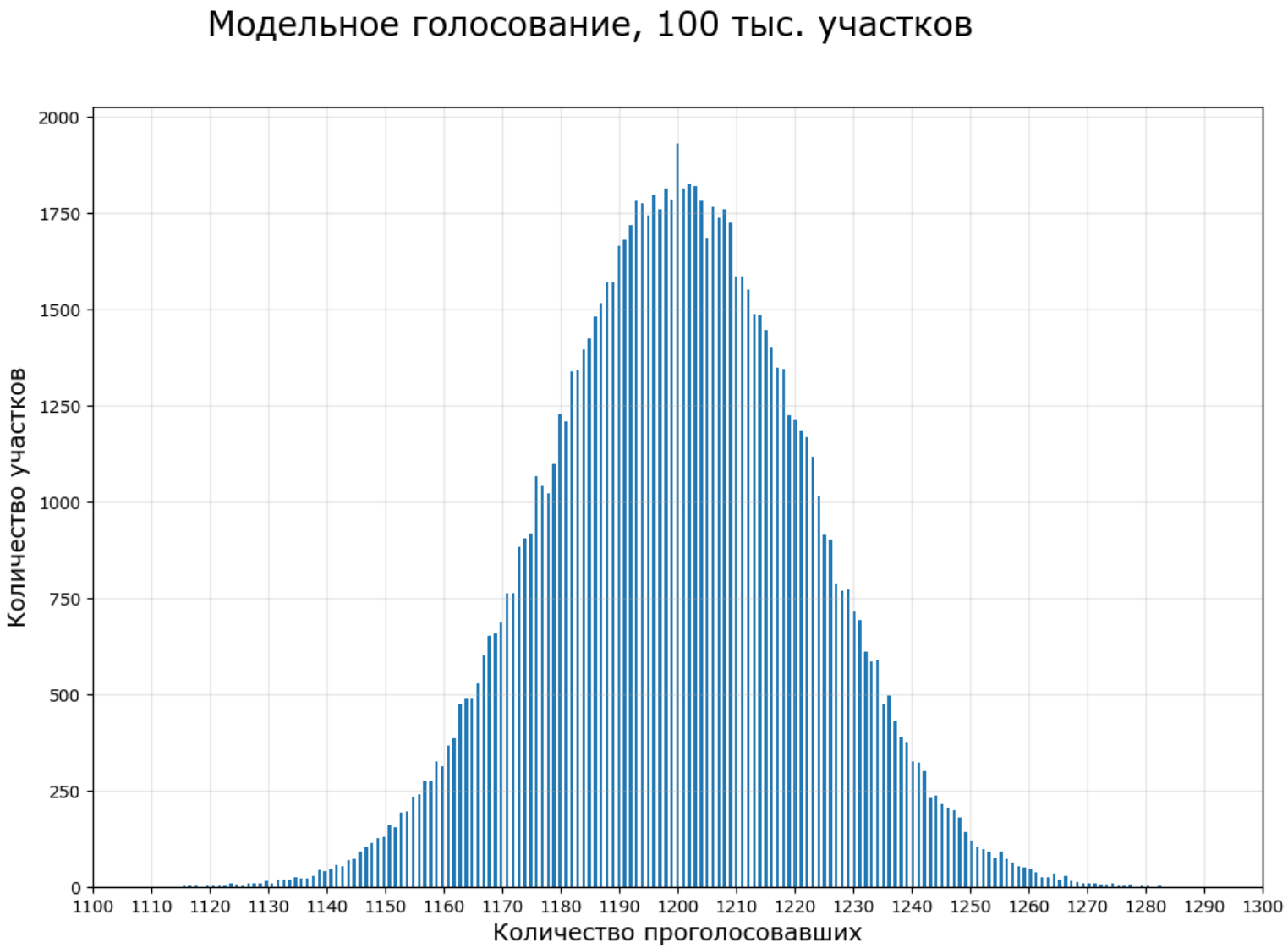
Модель избирательного участка

- Численное моделирование - 100 участков:

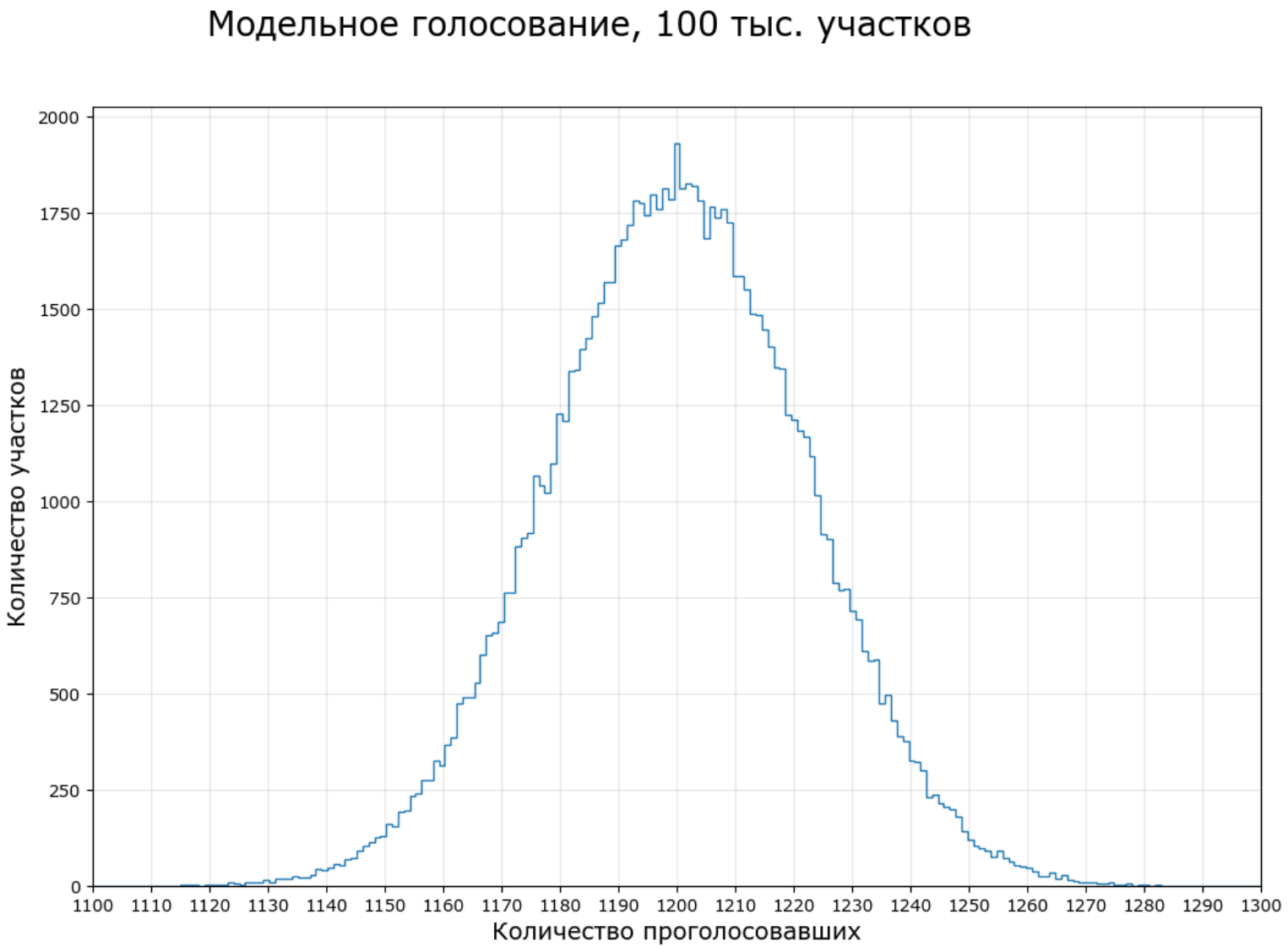
1182 1214 1162 1224 1211 1227 1196 1178 1226 1208
1162 1216 1197 1173 1230 1190 1190 1221 1198 1162
1227 1175 1199 1236 1177 1204 1160 1205 1203 1194
1216 1213 1180 1186 1199 1223 1176 1225 1183 1237
1218 1226 1257 1182 1197 1179 1212 1234 1193 1185
1206 1206 1242 1182 1194 1174 1204 1204 1192 1237
1202 1200 1180 1192 1173 1212 1173 1202 1169 1181
1175 1201 1231 1181 1220 1220 1179 1241 1184 1174
1133 1221 1204 1184 1190 1206 1194 1239 1192 1212
1186 1169 1198 1190 1189 1219 1198 1174 1199 1235

Модель избирательного участка

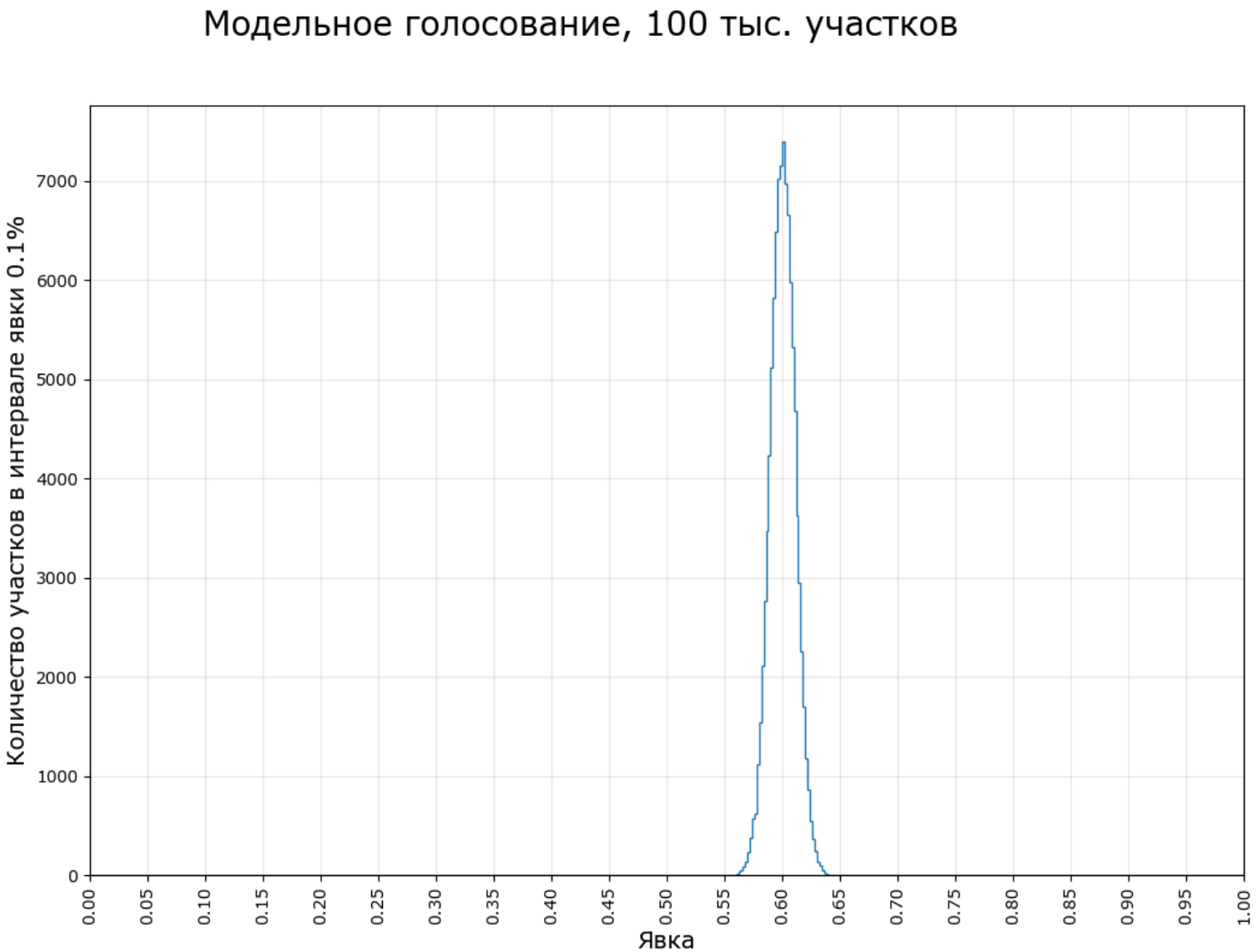
Модель избирательного участка



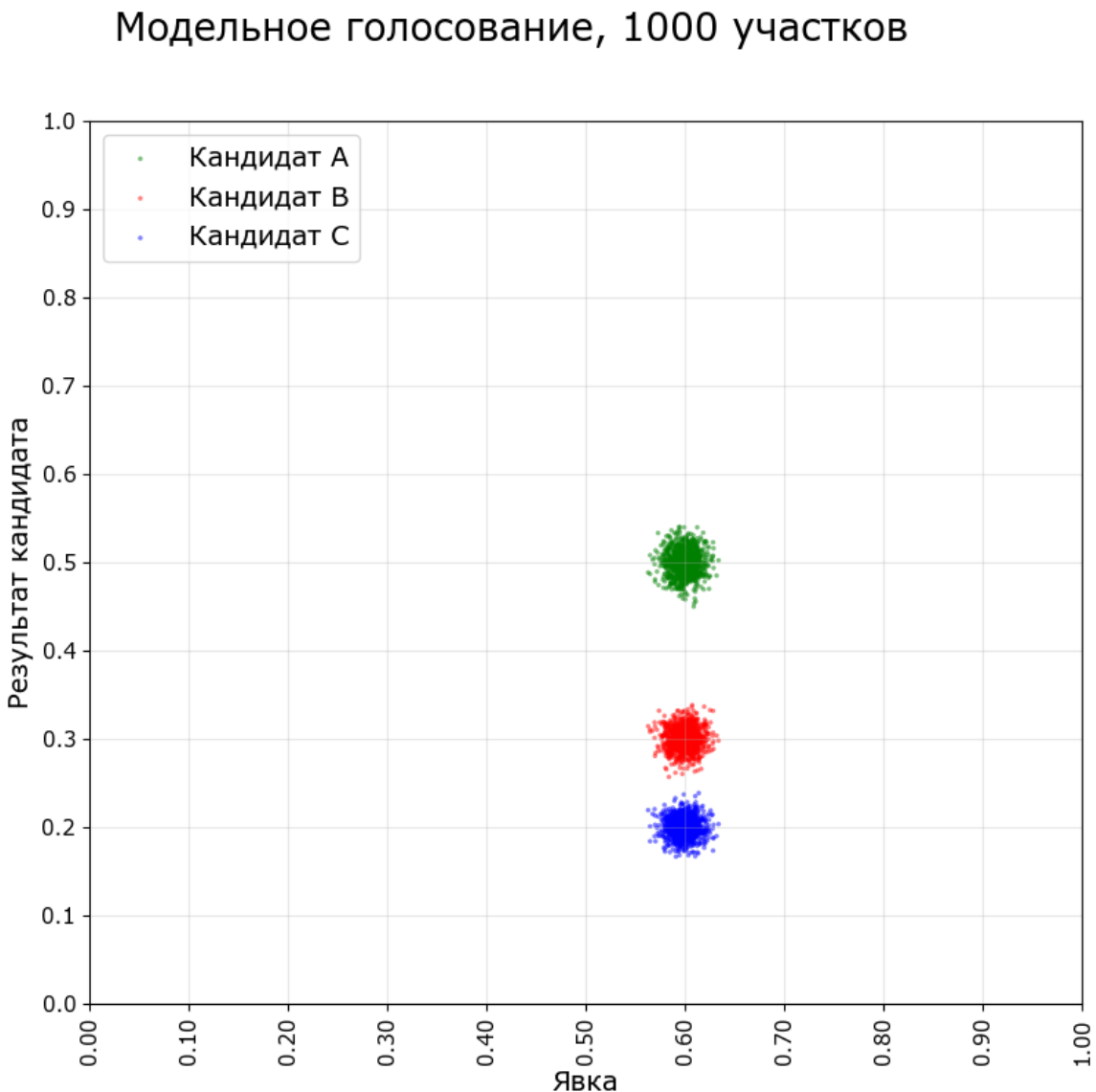
Модель избирательного участка



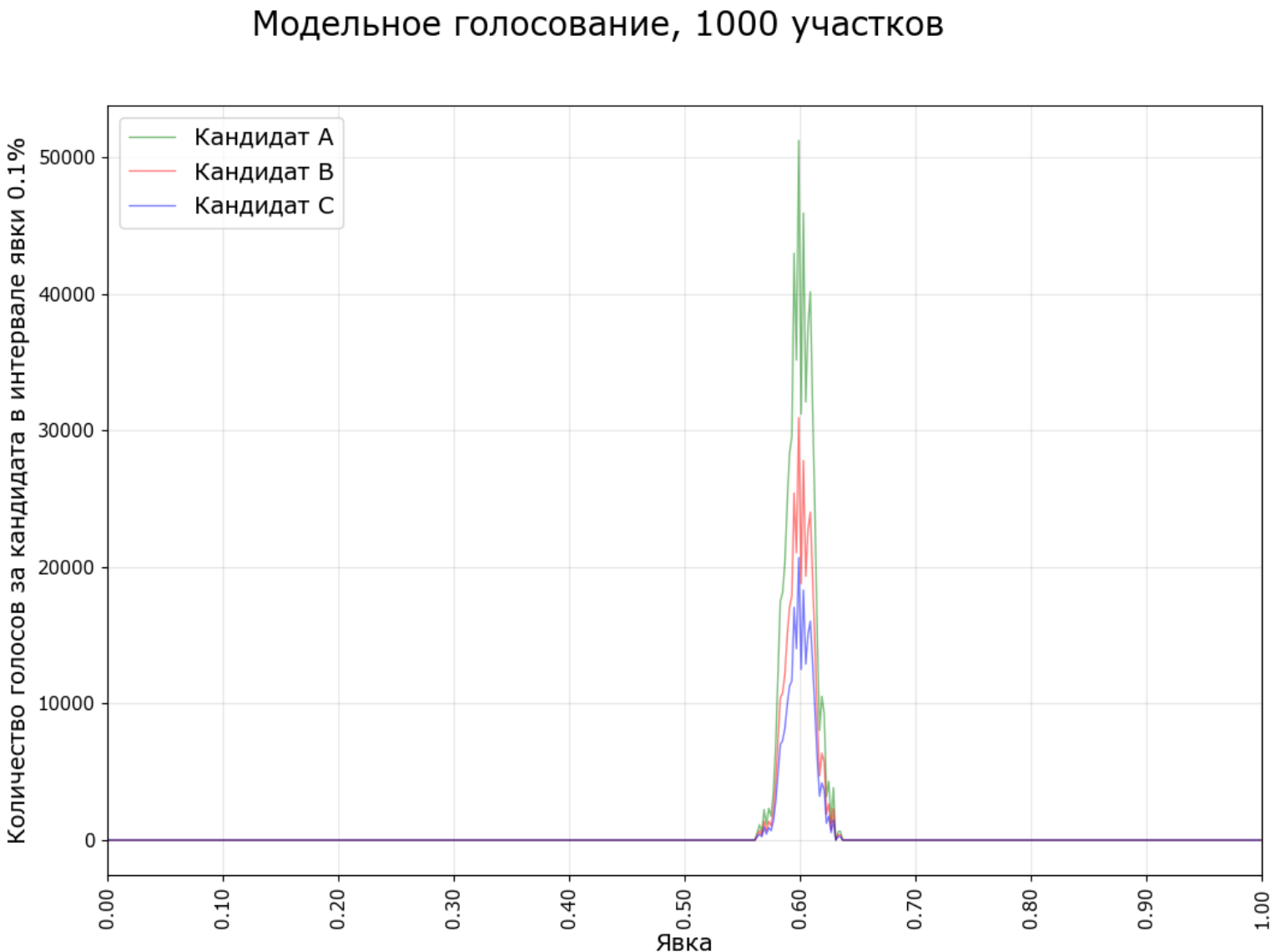
Модель избирательного участка



Модель избирательного участка



Модель избирательного участка

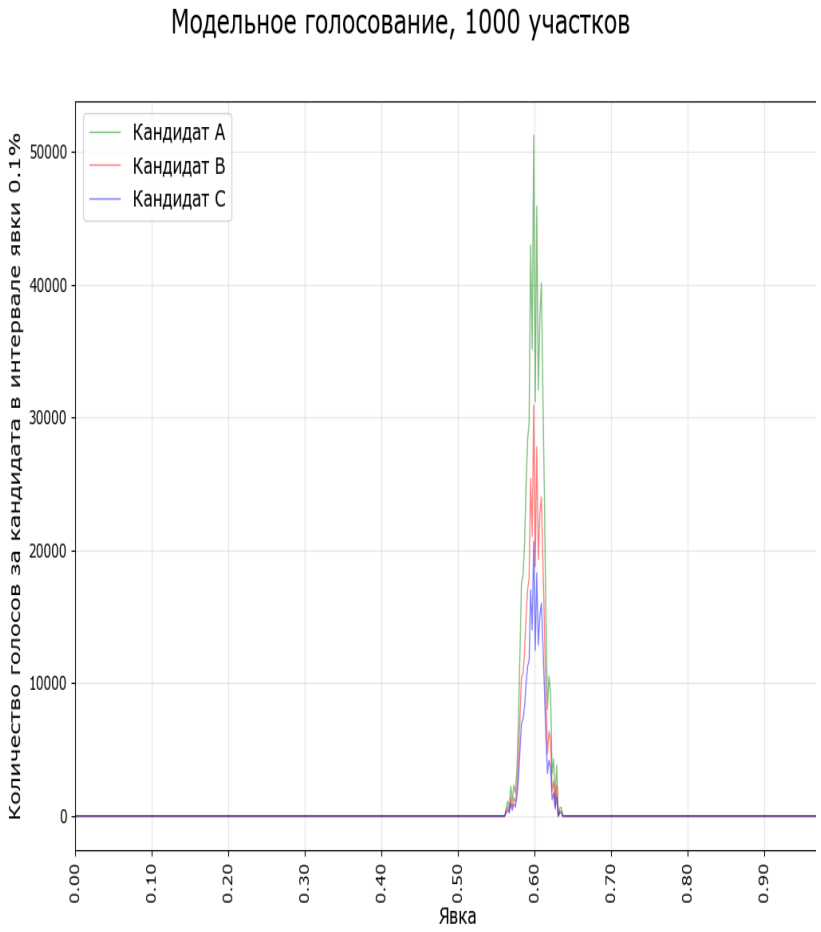


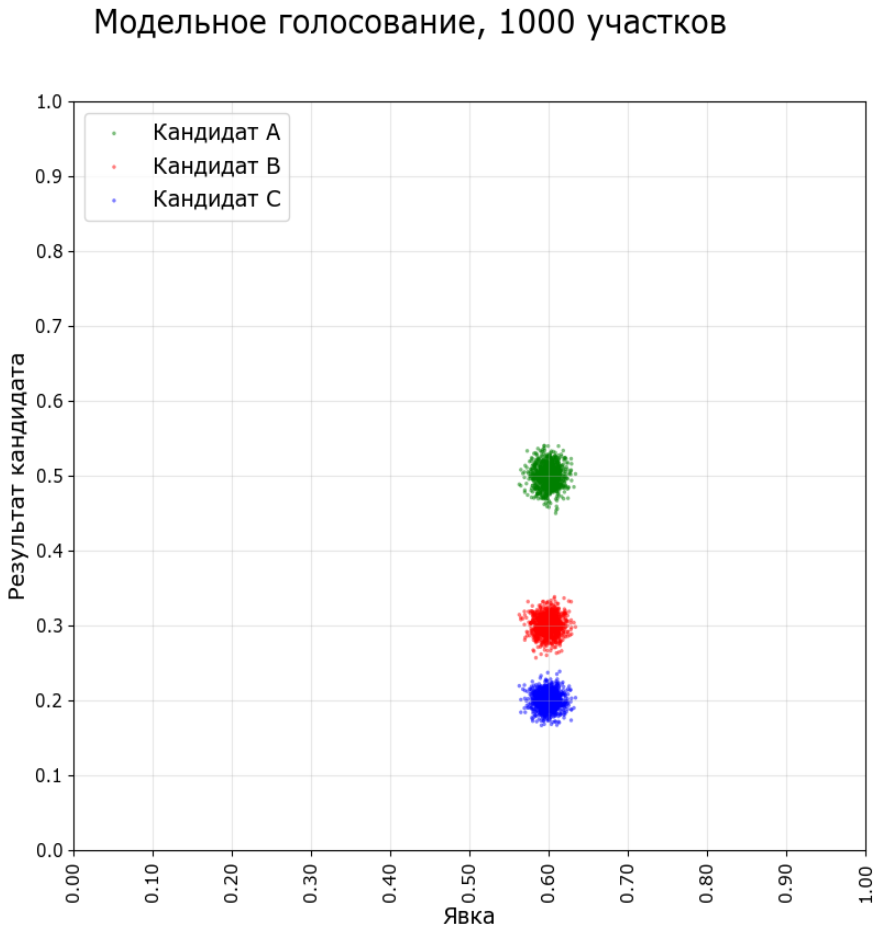
Важно:

- Даже одинаковые участки с одинаковыми по характеристикам избирателями при голосовании дают разные результаты
- Для типичного городского участка (2000 человек) при типичной явке 60% статистическое отклонение от средней явки составляет ~1%

Реальная ситуация:

- Участки неодинаковые по характеристикам
- Случайные факторы на уровне участка действуют в разные стороны от среднего – распределение явки размывается
- Ожидаемый разброс больше


Модельное голосование, 1000 участков
Кандидат A
Кандидат B
Кандидат C
Количество голосов за кандидата в интервале явки 0.1%
0
10000
20000
30000
40000
50000
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
Явка

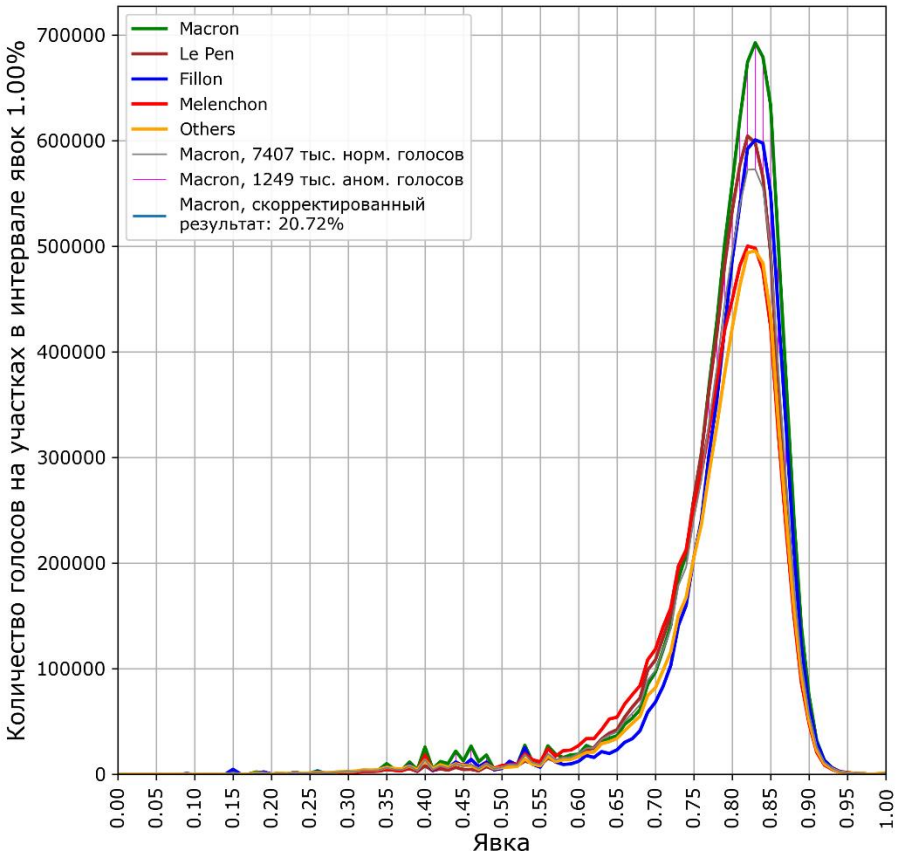

Модельное голосование, 1000 участков
Кандидат A
Кандидат B
Кандидат C
Результат кандидата
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
Явка

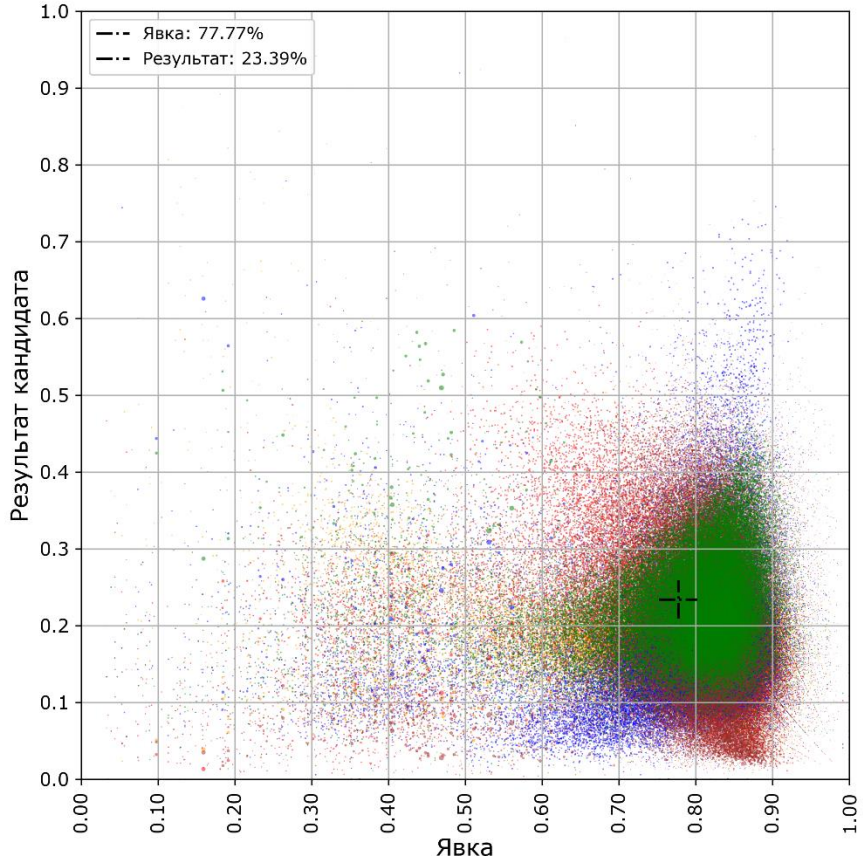
Почему сохраняется колокол?

- На явку действуют много разнонаправленных случайных факторов
- Усредненные воздействия этих факторов вносят вклад в среднее значение явки
- Отклонения от среднего с большей вероятностью разнонаправлены, чем сонаправлены

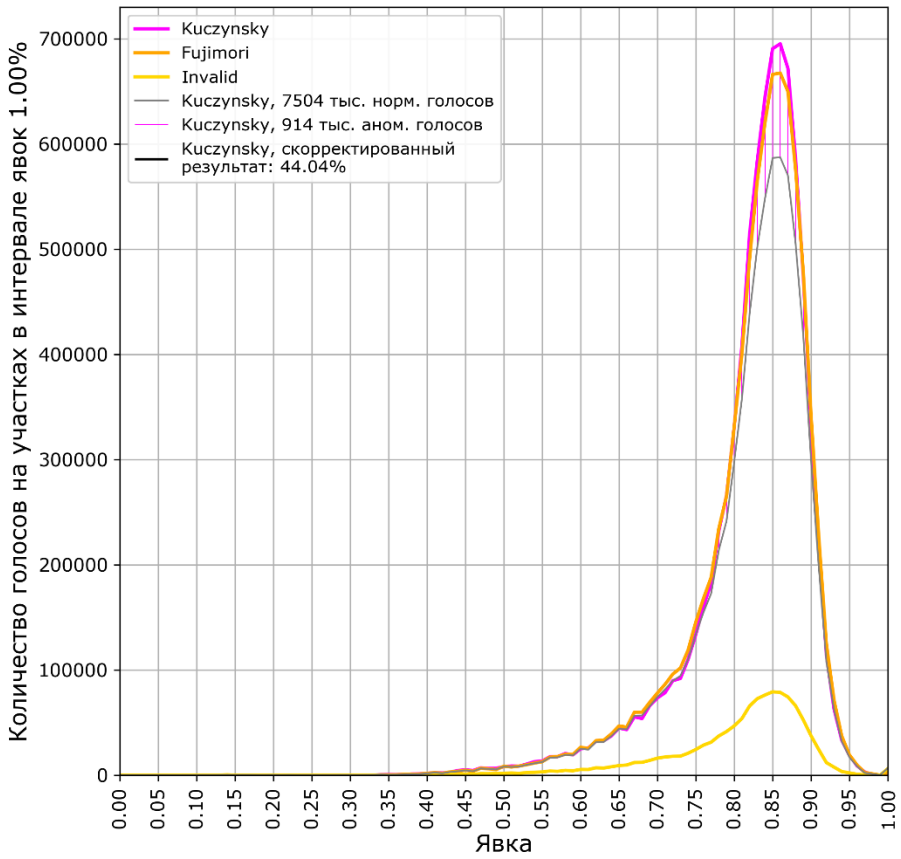
- Чтобы разрушить колокол, нужен единичный сильный фактор, превосходящий по воздействию суммарное действие остальных

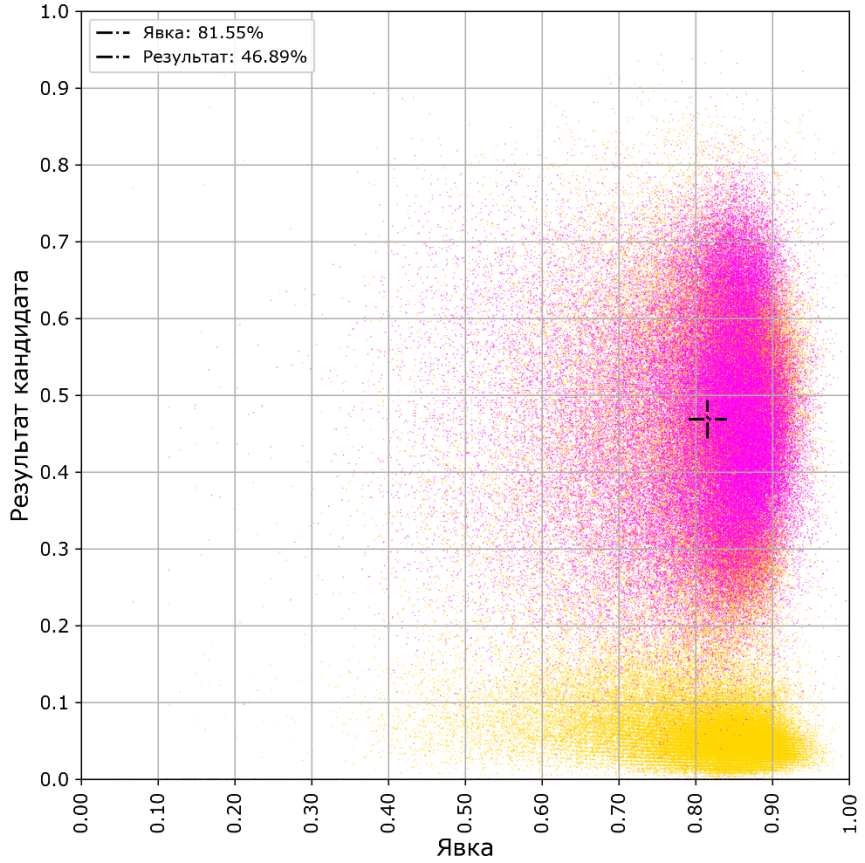
Entire country
France President 2017 Round_1



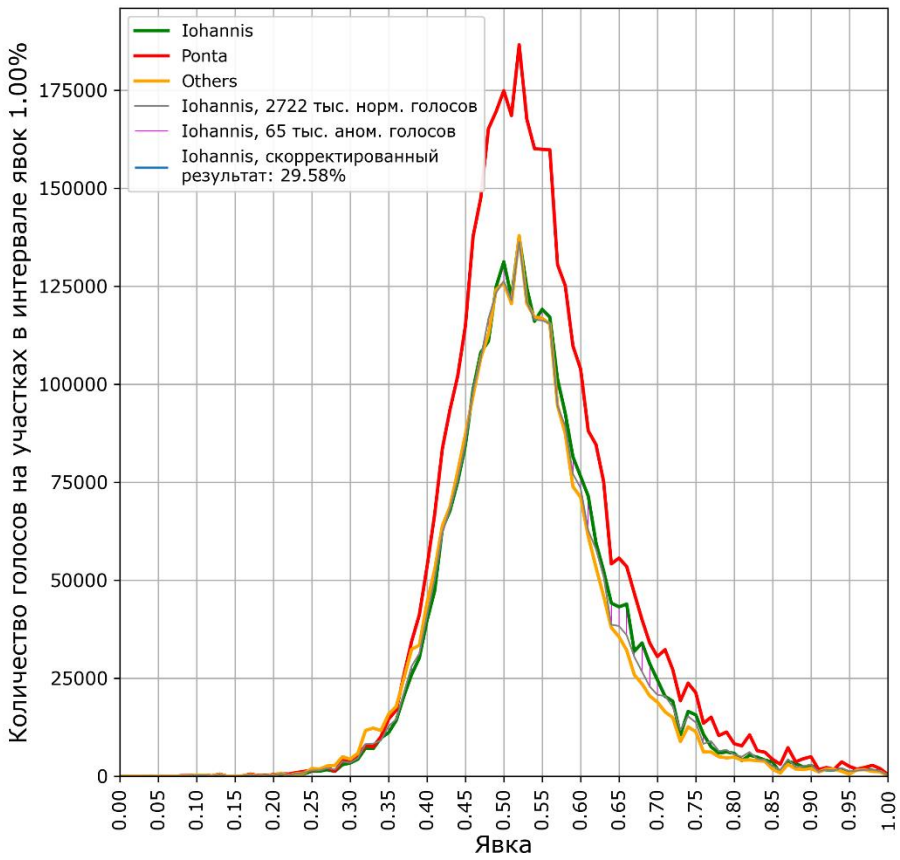


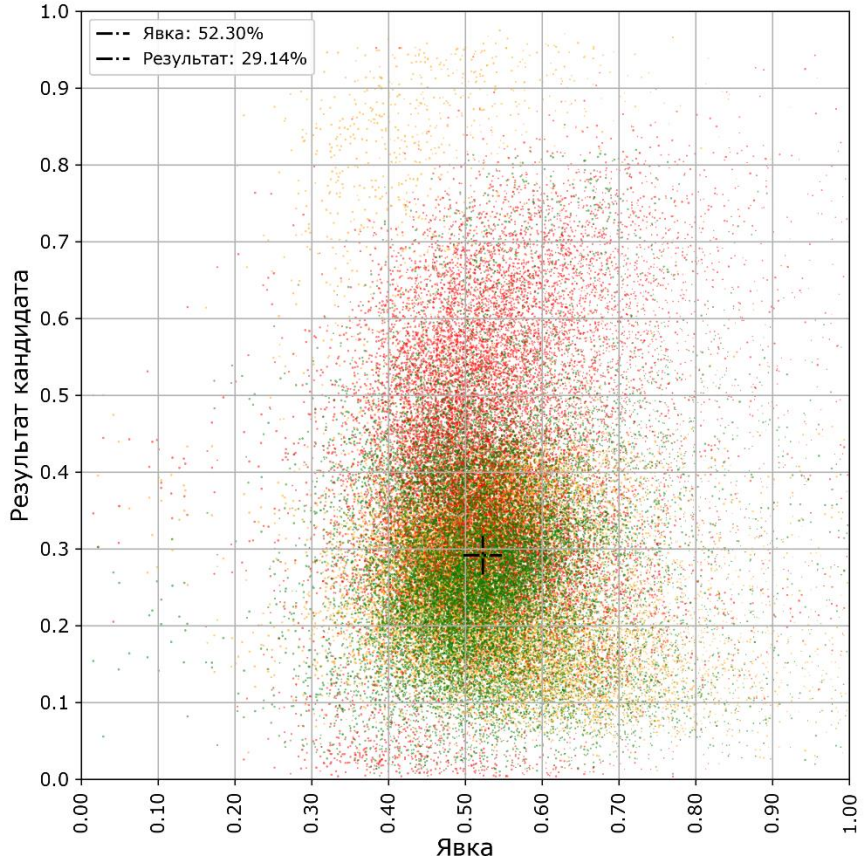
Whole country
Peru precidenciales Tour 2





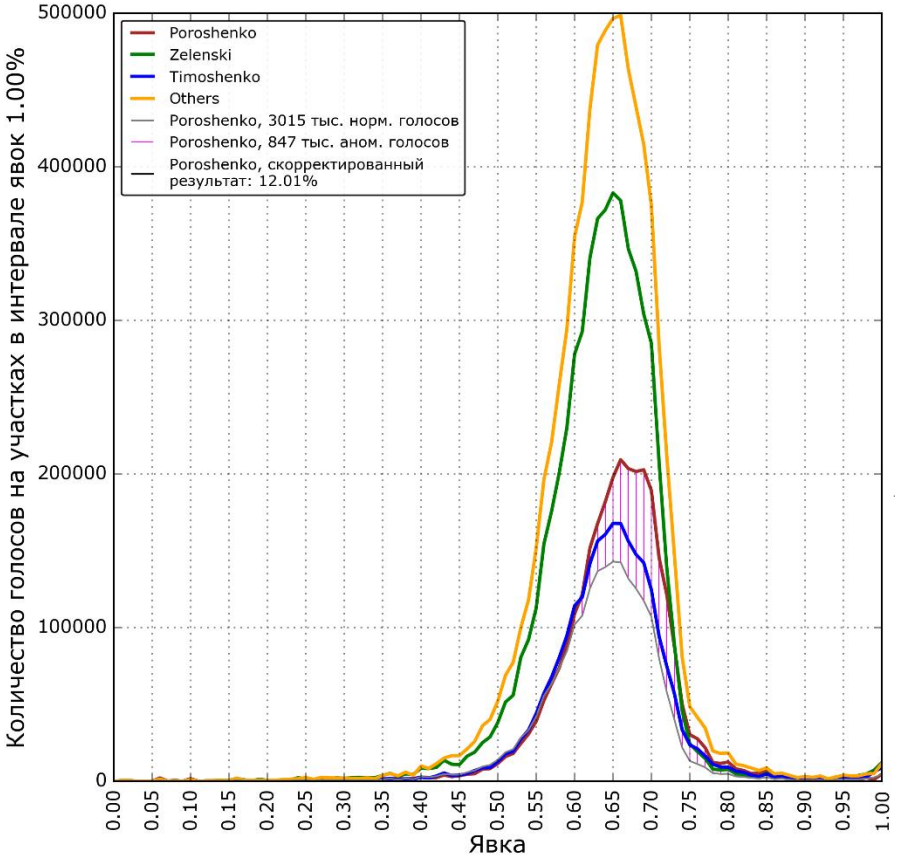
Entire country
Romania President 2014 Tour_1

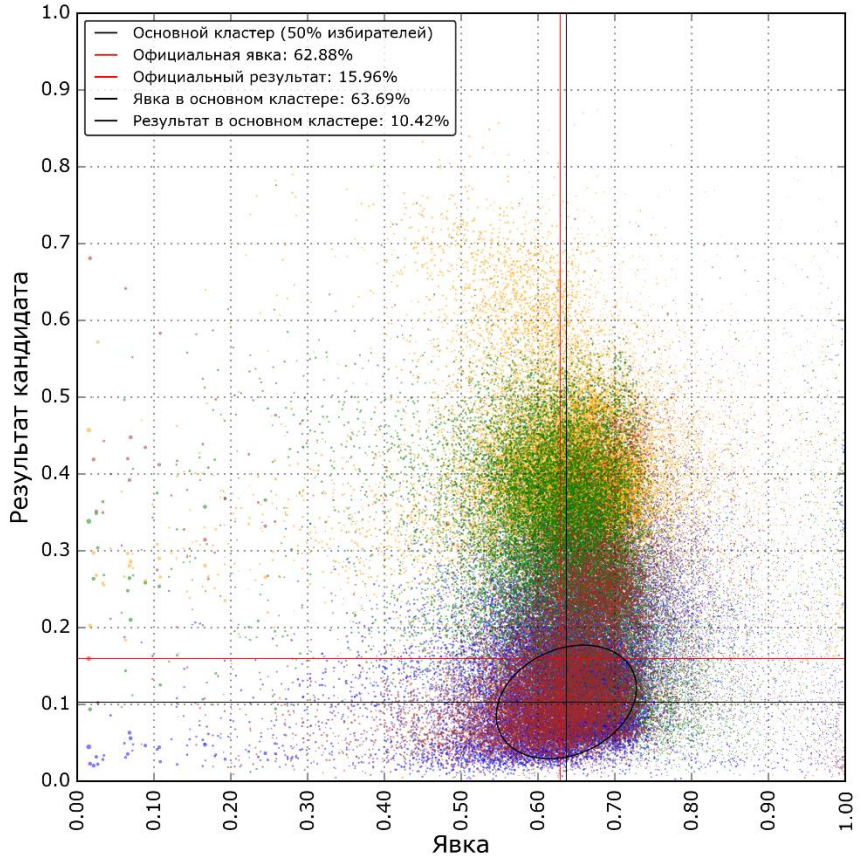




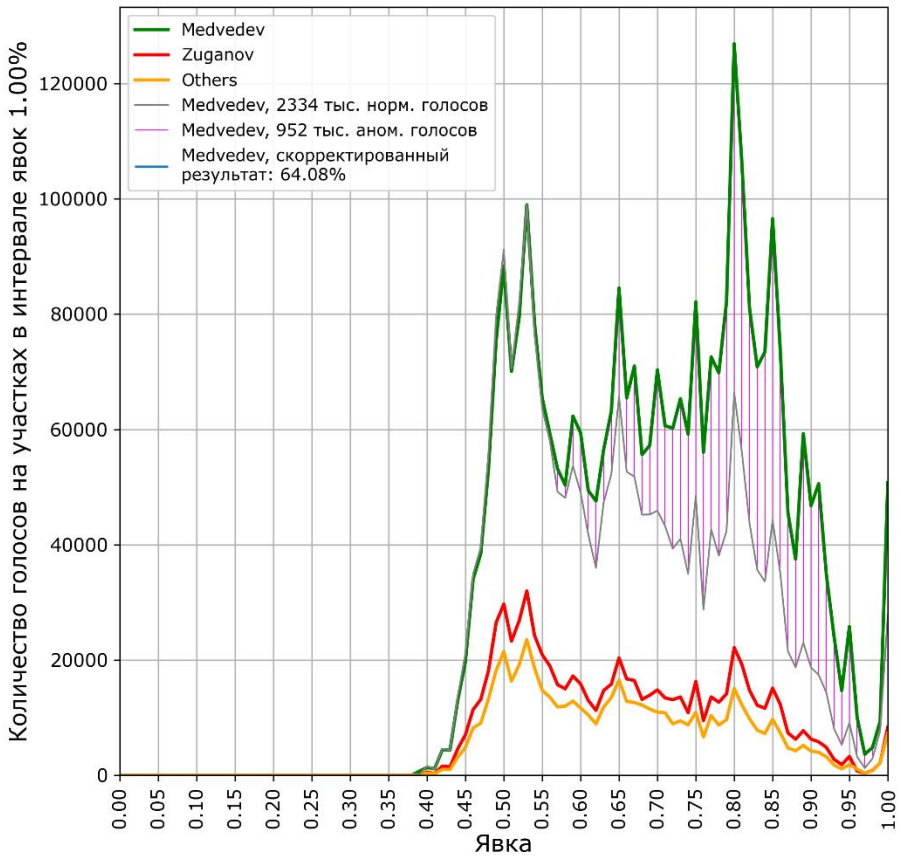
Entire region

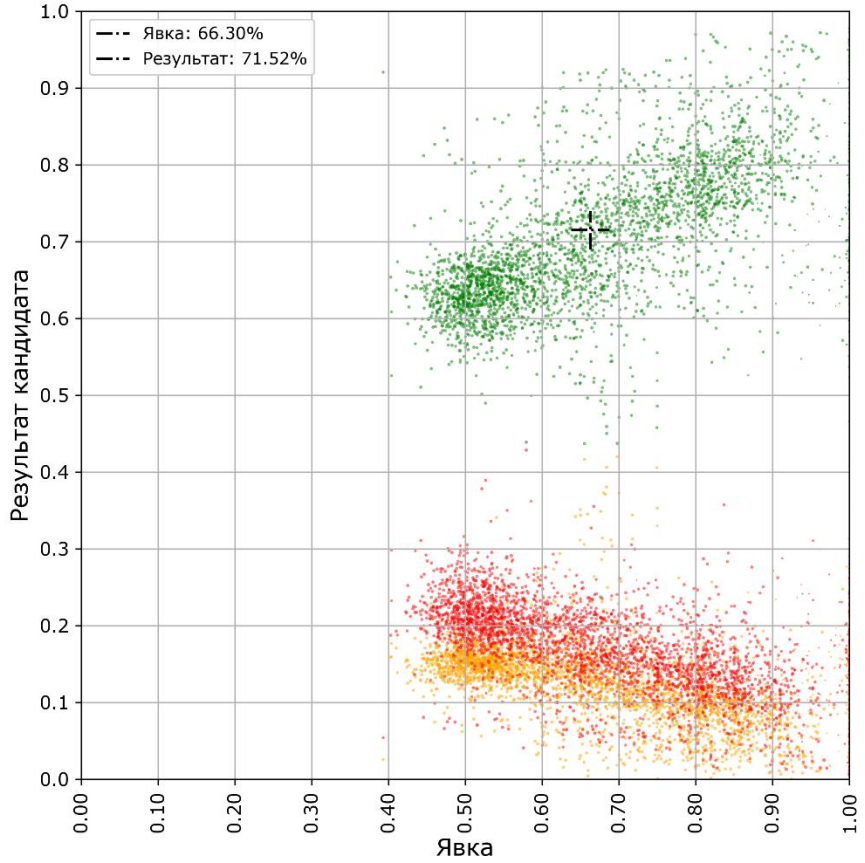
Украина – 2019



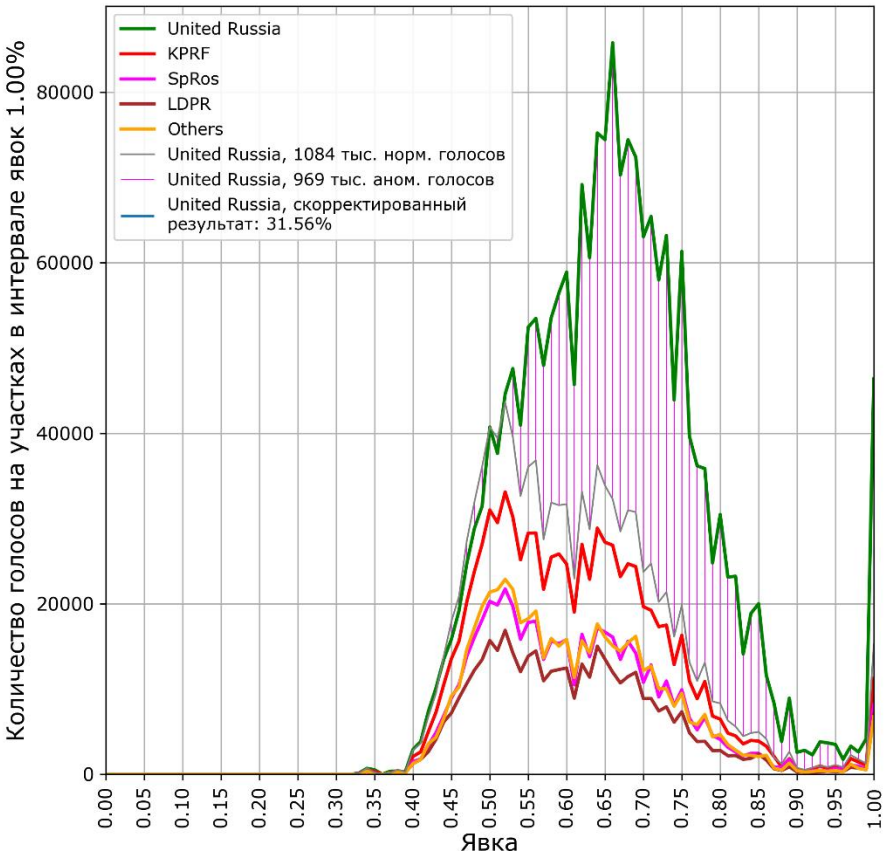


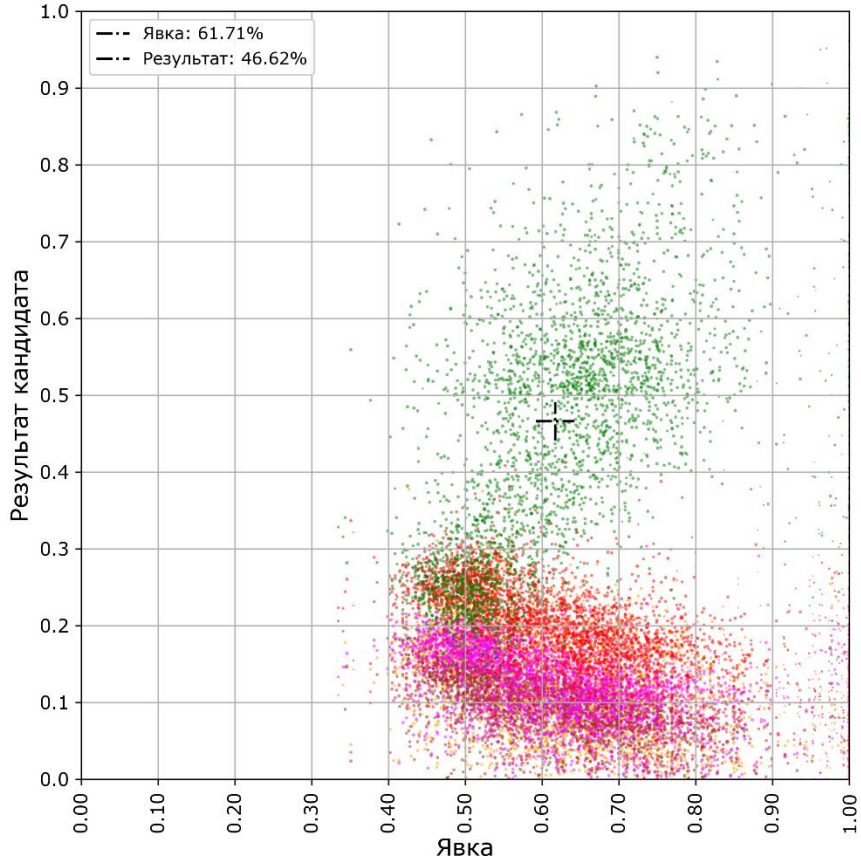
Город Москва
Russia President election 2008



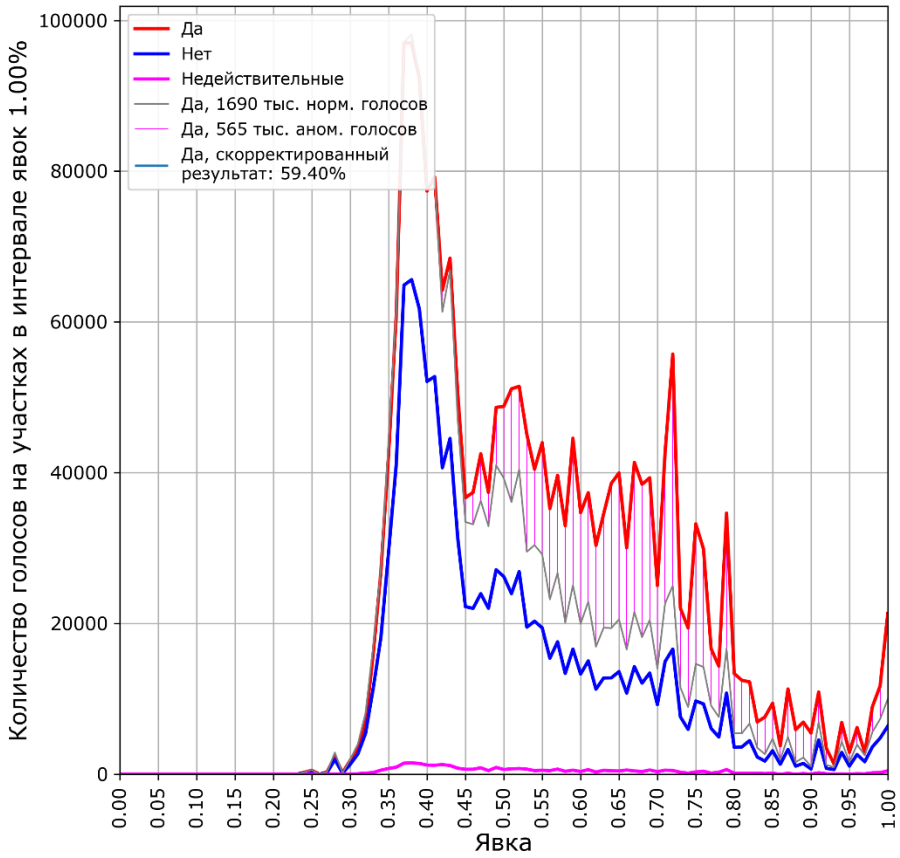


Город Москва
Russia Duma election 2011

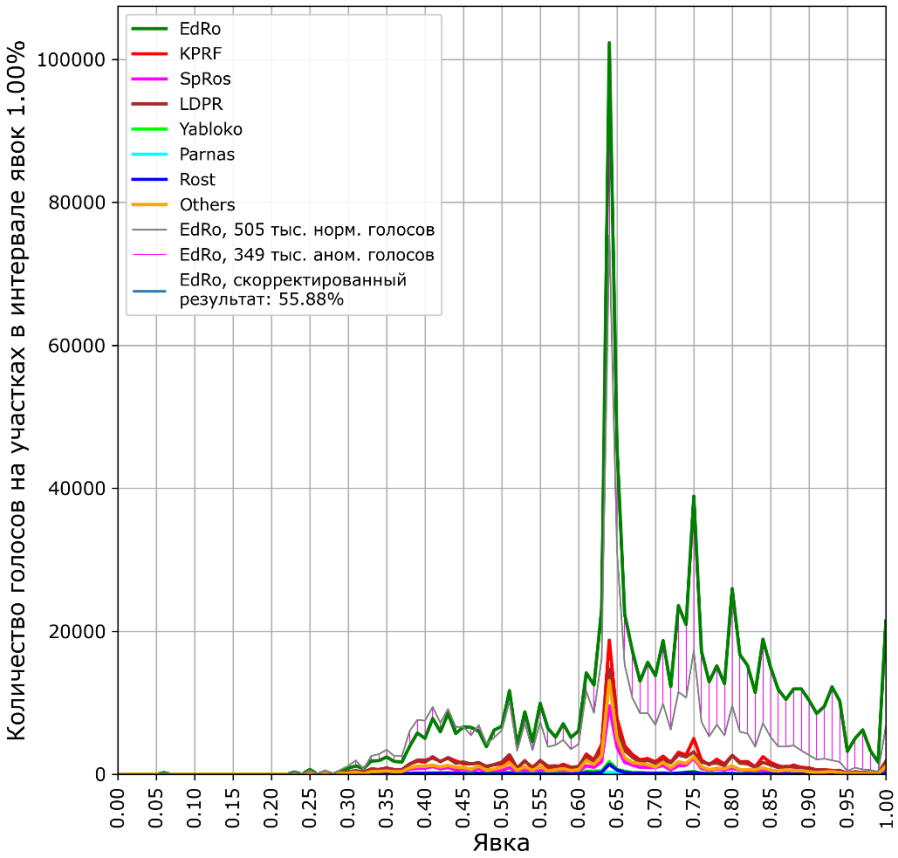


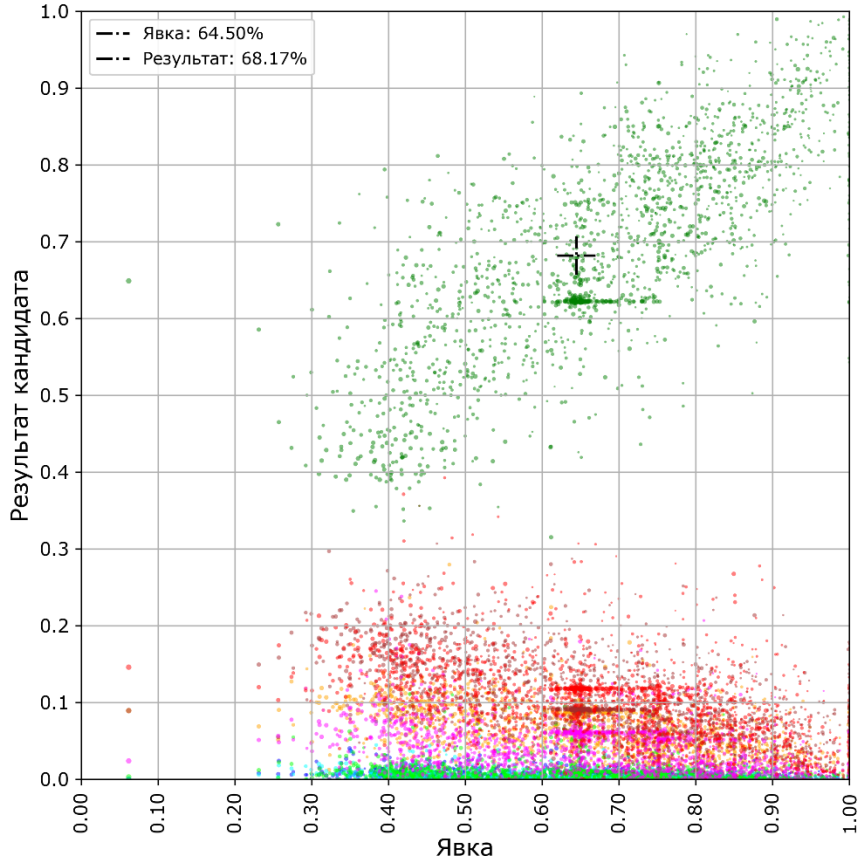


город Москва
Референдум 2020



Саратовская область
Russia Duma election 2016





Саратов-2016

	Недейств	Явка	Родина	КР	Пенс	ЕР	Зелен	ГП	ЛДПР	Парнас	Рост	ГС	Яблоко	КПРФ	ПР	СР
УИК №69	**0,0%**	**73,1%**	**2,3%**	**3,3%**	**0,3%**	**62,3%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №70	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,2%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №71	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №72	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №73	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №74	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,2%**
УИК №75	**0,0%**	**67,2%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №76	**0,8%**	**70,0%**	**2,3%**	**3,3%**	**0,3%**	**61,8%**	**0,7%**	**0,2%**	**9,0%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,7%**	**1,5%**	**6,1%**
УИК №77	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,0%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №78	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №79	**0,0%**	**67,0%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №80	**0,2%**	**69,4%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,3%**	**6,1%**
УИК №81	**0,1%**	**69,9%**	**2,7%**	**3,3%**	**0,3%**	**62,2%**	**0,6%**	**0,2%**	**9,1%**	**0,1%**	**0,8%**	**0,5%**	**1,0%**	**11,7%**	**1,5%**	**6,1%**
УИК №82	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,4%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,7%**	**1,5%**	**6,2%**
УИК №83	**0,0%**	**55,0%**	**2,7%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,1%**	**9,1%**	**0,1%**	**0,8%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №84	**0,0%**	**65,0%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №85	**0,0%**	**64,8%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №86	**0,0%**	**69,7%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №87	**0,0%**	**64,3%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**
УИК №88	**0,0%**	**75,0%**	**2,3%**	**3,3%**	**0,3%**	**62,2%**	**0,7%**	**0,2%**	**9,1%**	**0,1%**	**0,9%**	**0,5%**	**1,0%**	**11,8%**	**1,5%**	**6,1%**

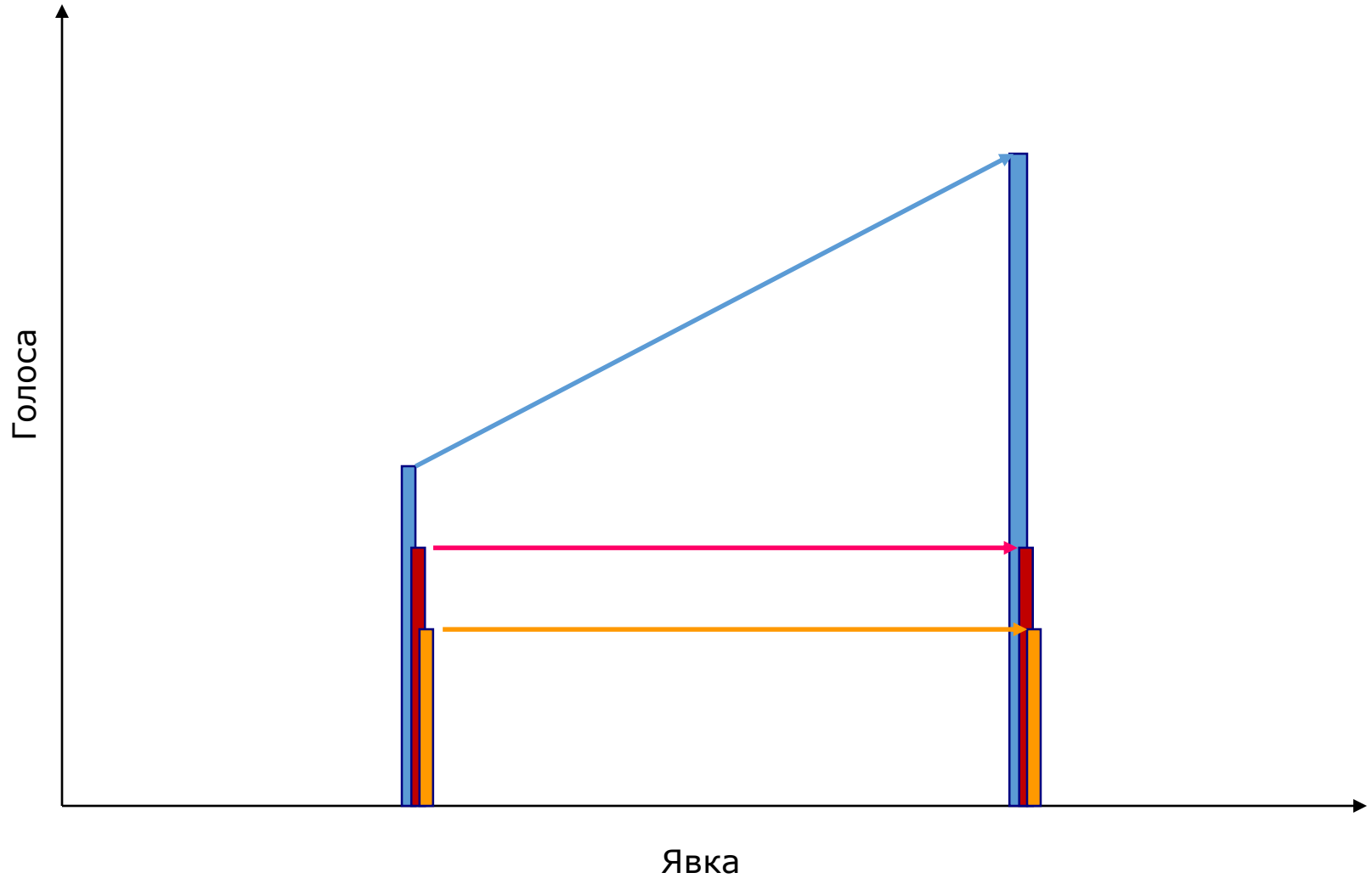
161.ru Вброс на УИК №1749 Ростов. 18 сентября 2016 года. 161.ru

Виды фальсификаций на участке

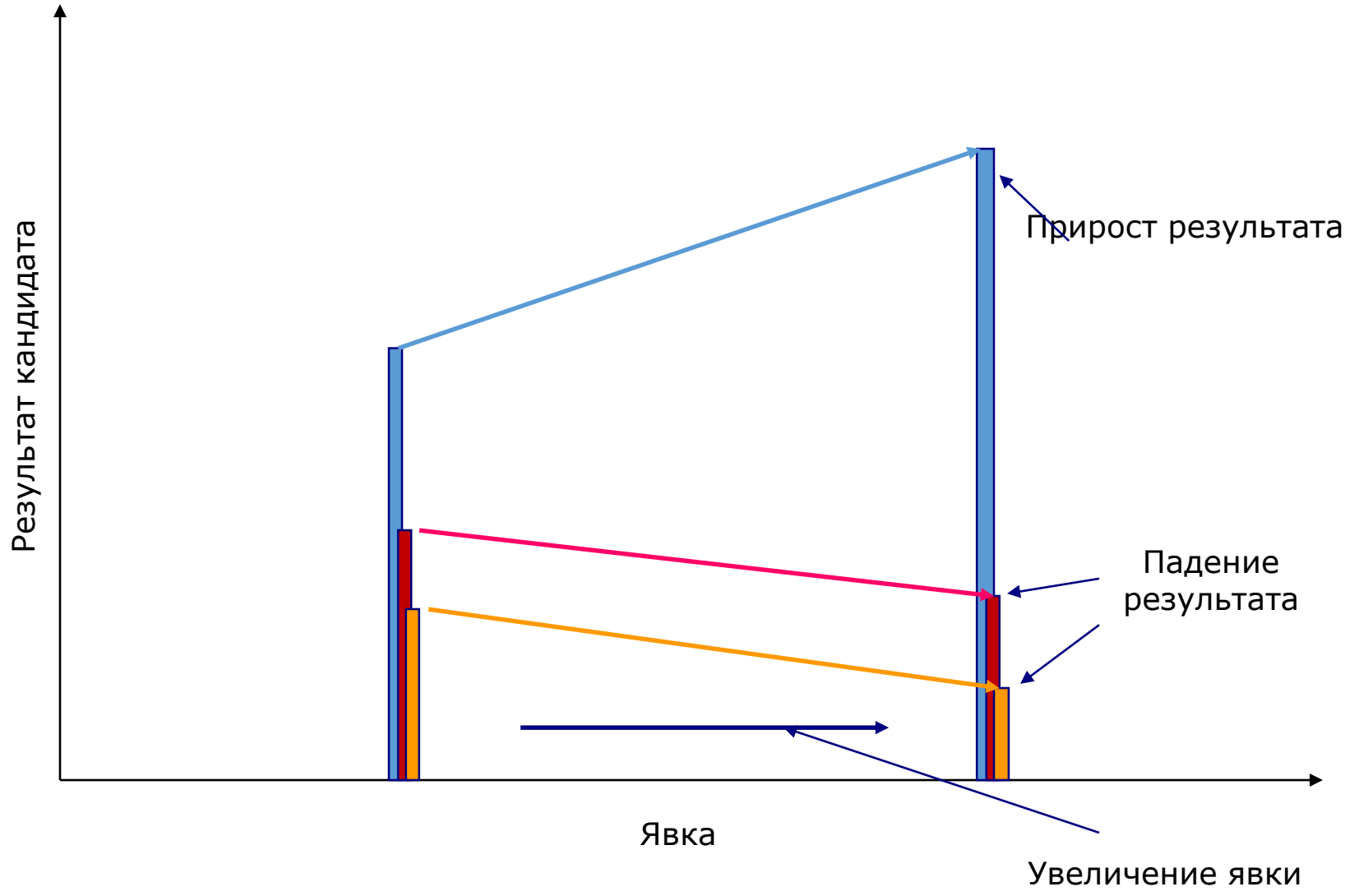
- Вброс бюллетеней/приписывание голосов
- Переброс бюллетеней от кандидата к кандидату (вариант: обмен голосов)
- Рисование голосов

- Вброс – самый технически «дешевый» (в случае КОИБ – основной) вариант

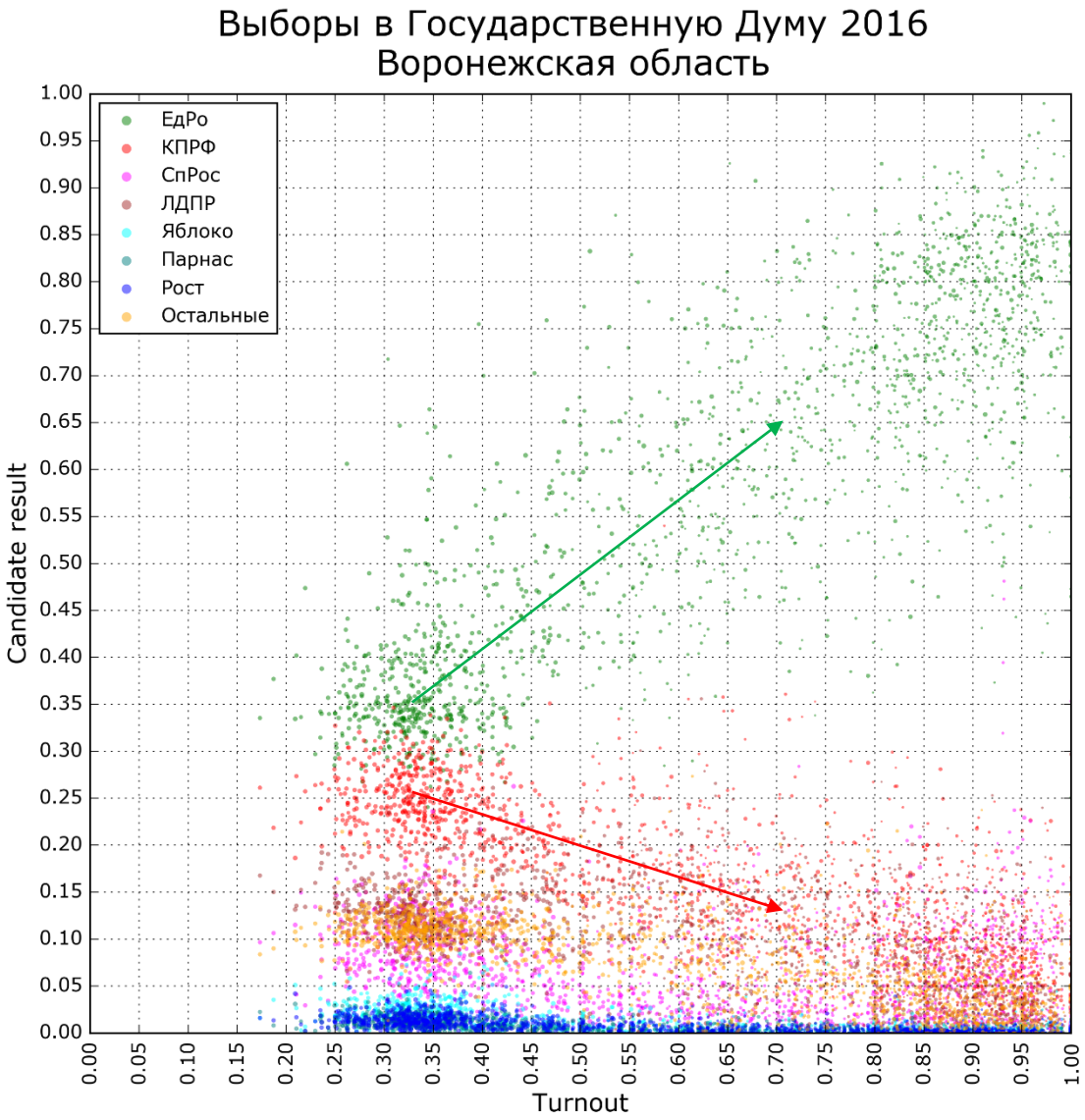
Добавление голосов за кандидата



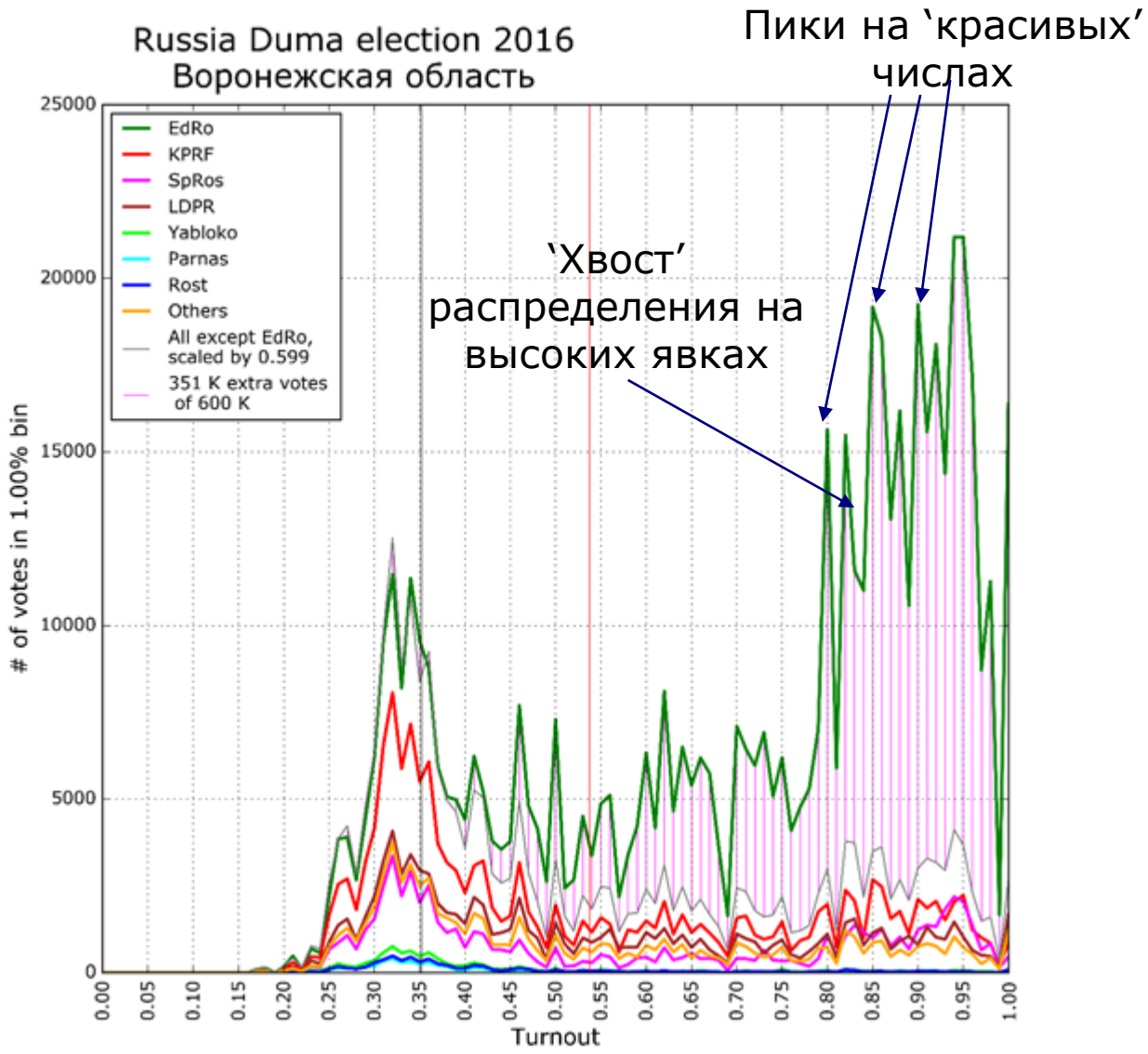
Добавление голосов за кандидата



Добавление голосов за кандидата



Добавление голосов за кандидата



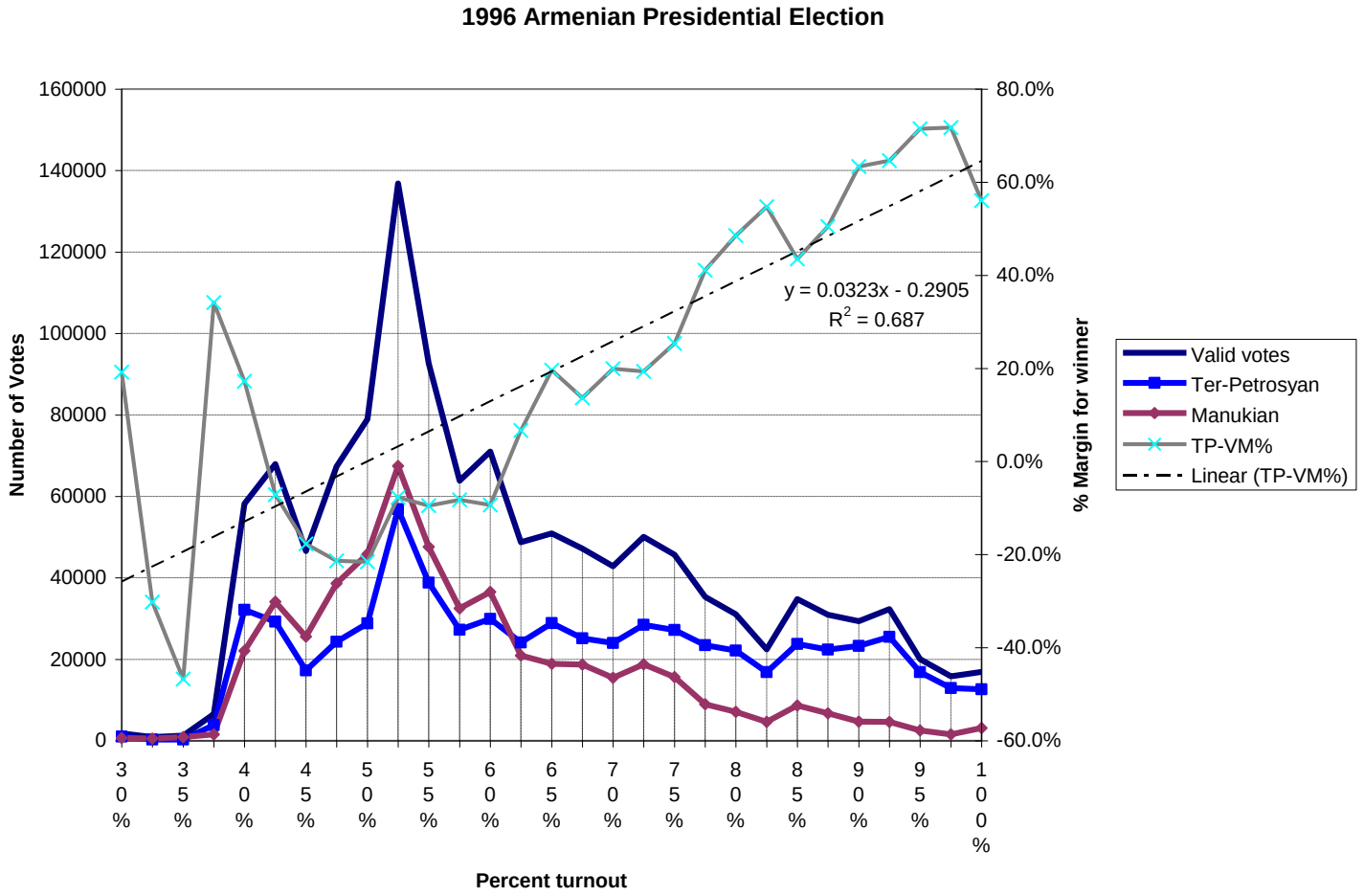
Вычисление вброса

- Цель: оценка реального результата в присутствии фальсифицированных данных
- Подходы: методы робастной (устойчивой к наличию отклоняющихся точек) статистики

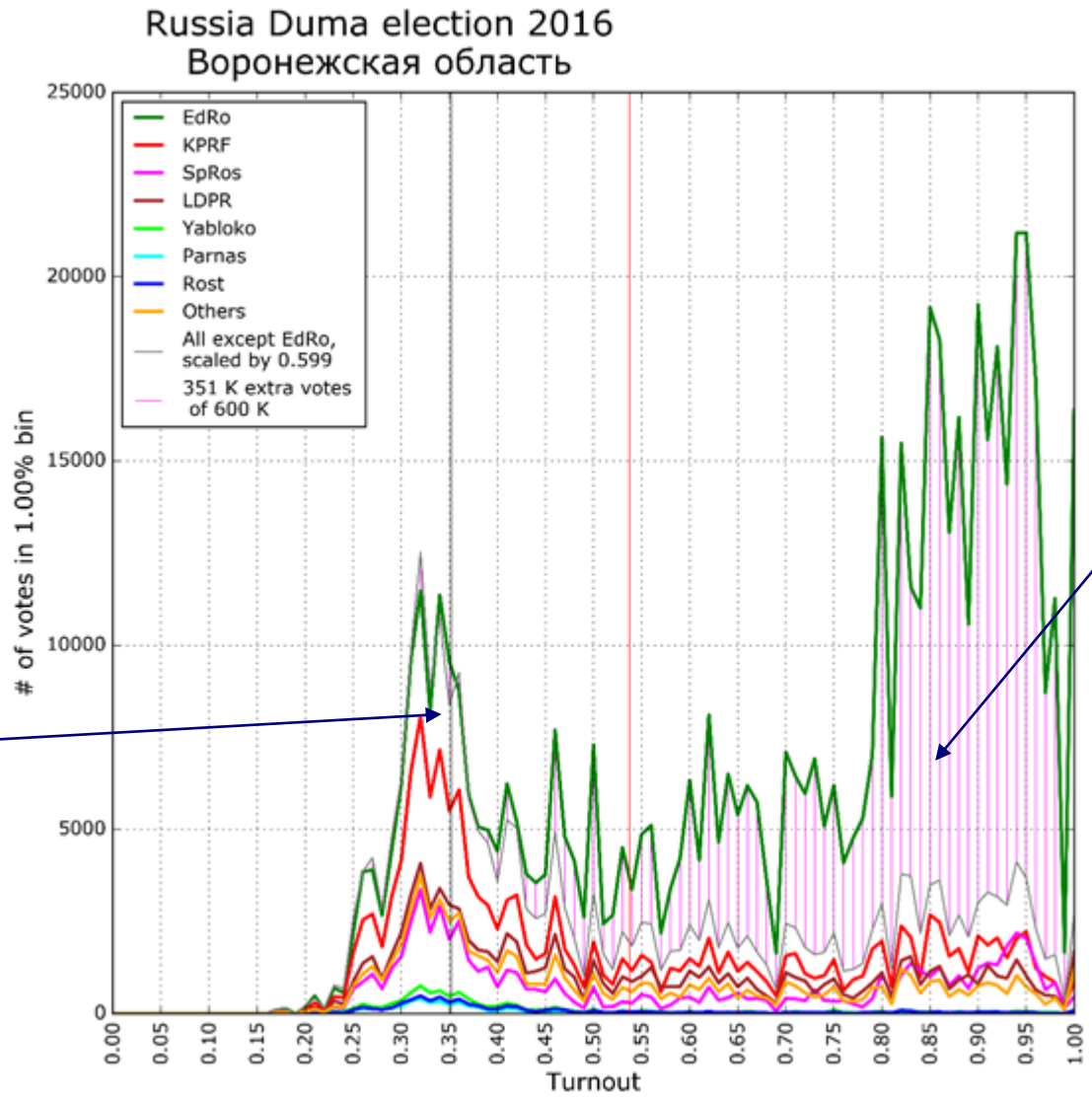
- В чем смысл?
  - Юридически имеют смысл только официальные результаты
- Уровень фальсификаций – индекс состояния политической системы

Вычисление вброса

- J.B.Kiesling (1996)



Вычисление вброса



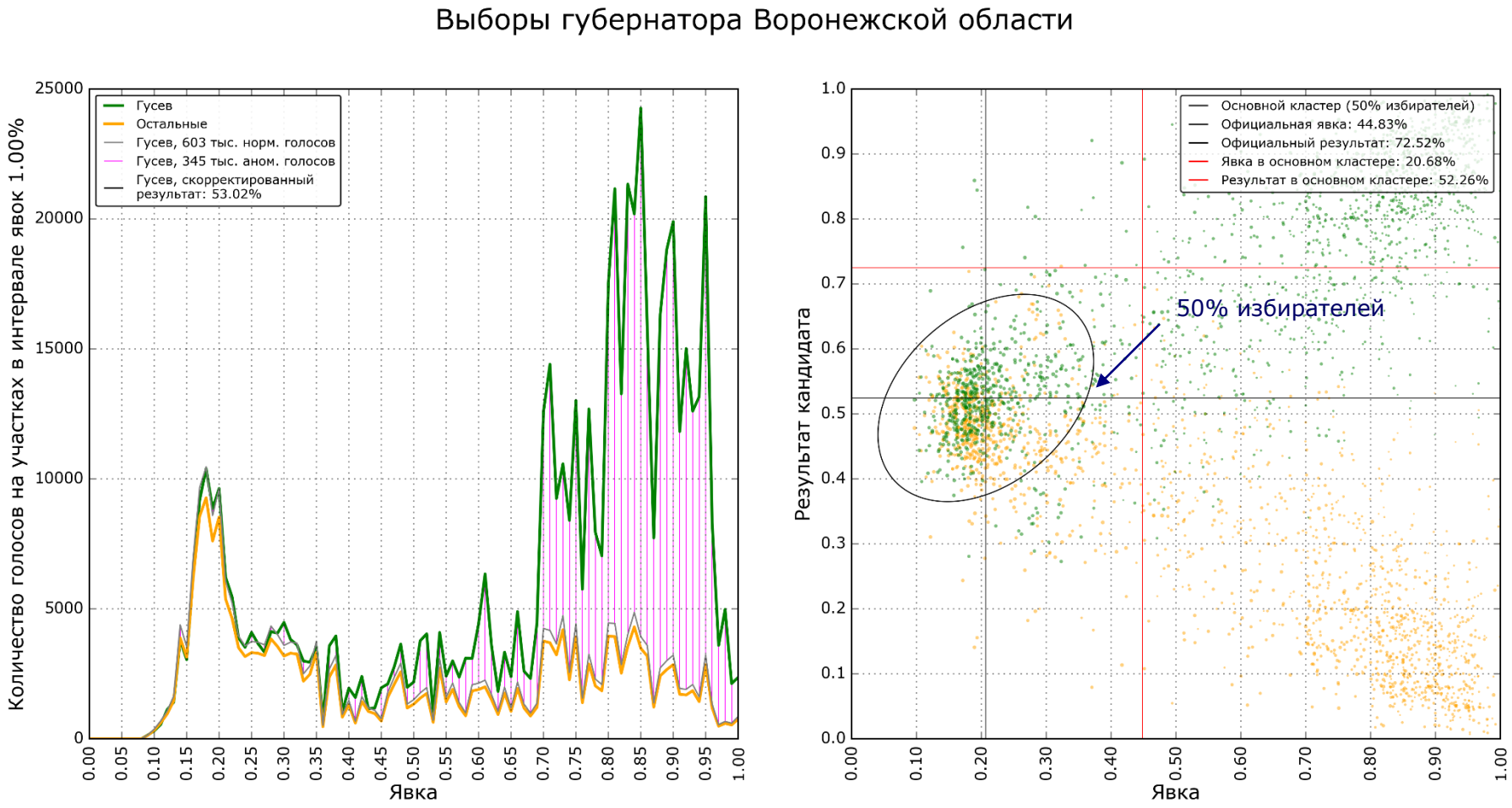
Начальные участки гистограмм совмещаются

Для чистого вброса/добавления голосов площадь заштрихованной области = количество вброшенных голосов

Вычисление вброса

- Процедура вычисления неоднозначна, как любая процедура робастной статистики, и зависит от выбранного алгоритма совмещения гистограмм
- При наличии переброса голосов между кандидатами/рисования результат становится оценочным, но может по-прежнему рассматриваться как индекс фальсификации

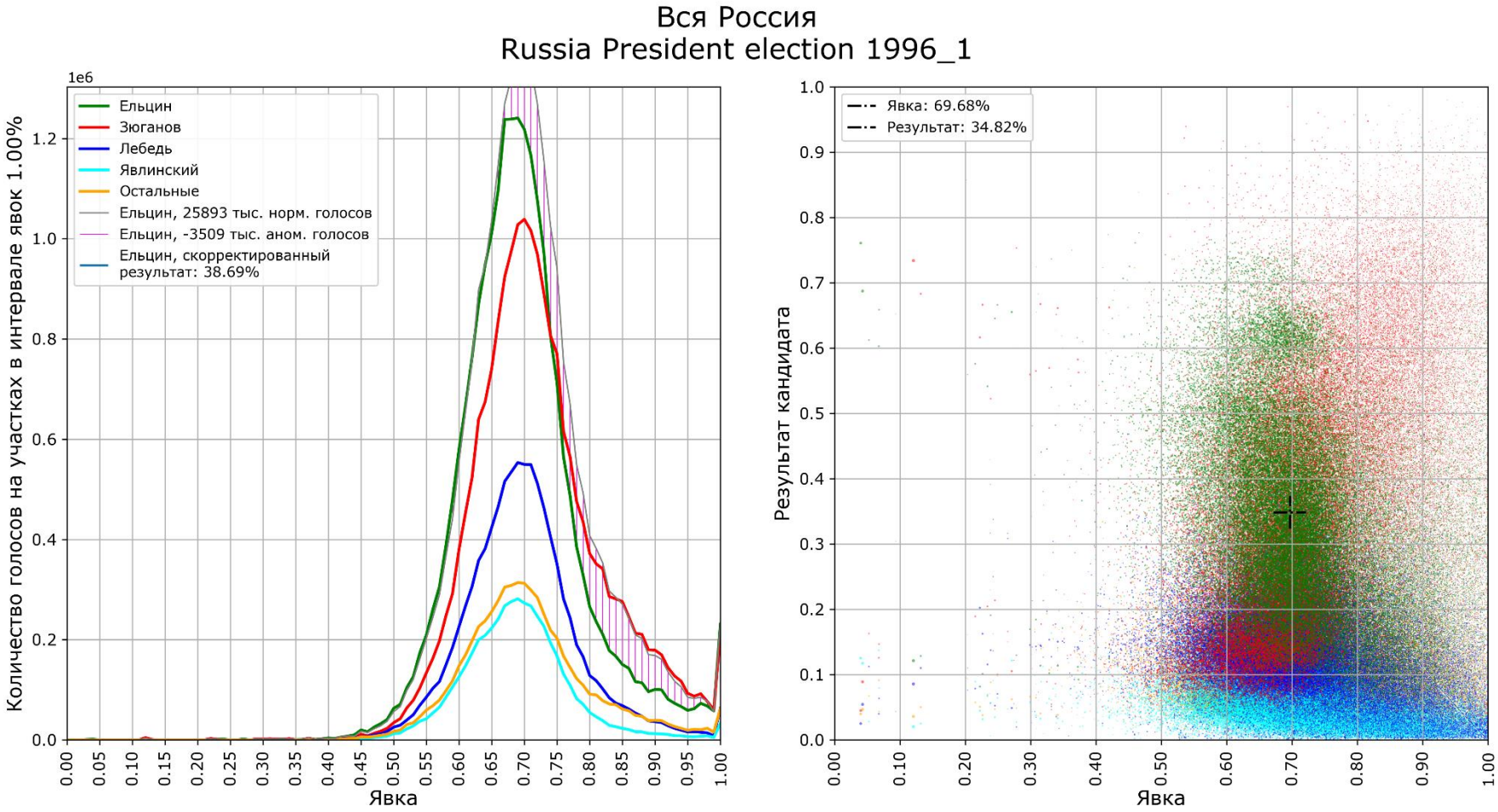
Альтернативная оценка: основной кластер диаграммы рассеяния



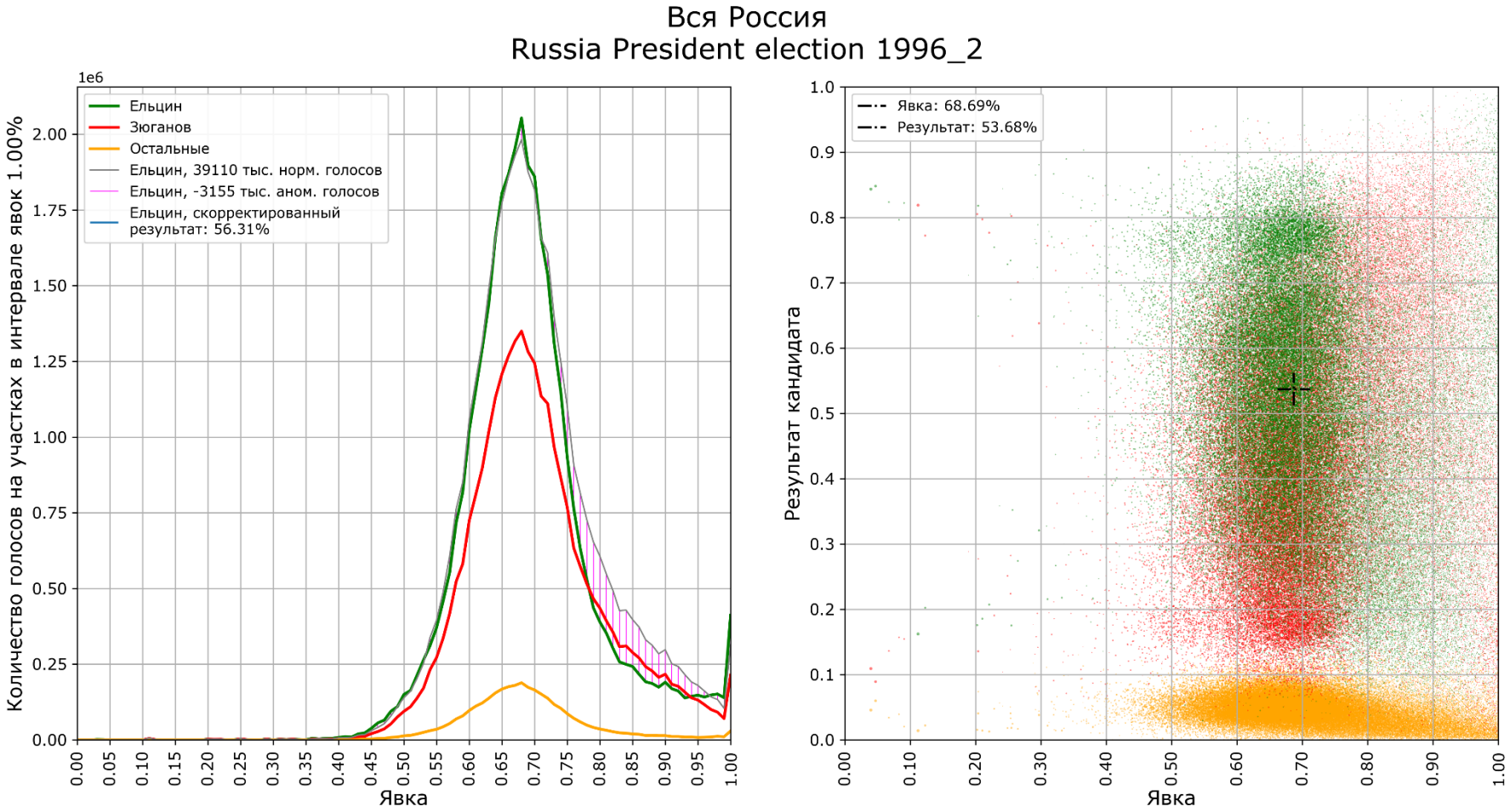
Альтернативная оценка: основной кластер диаграммы рассеяния

- Предполагается, что основной кластер один
  - Бывают регионы с двумя кластерами (Самарская, Вологодская область)
- При сильных фальсификациях основной кластер сам состоит из фальсифицированных данных
  - Проблема для сильно фальсифицированных регионов
  - В 2020 году – проблема для всей России

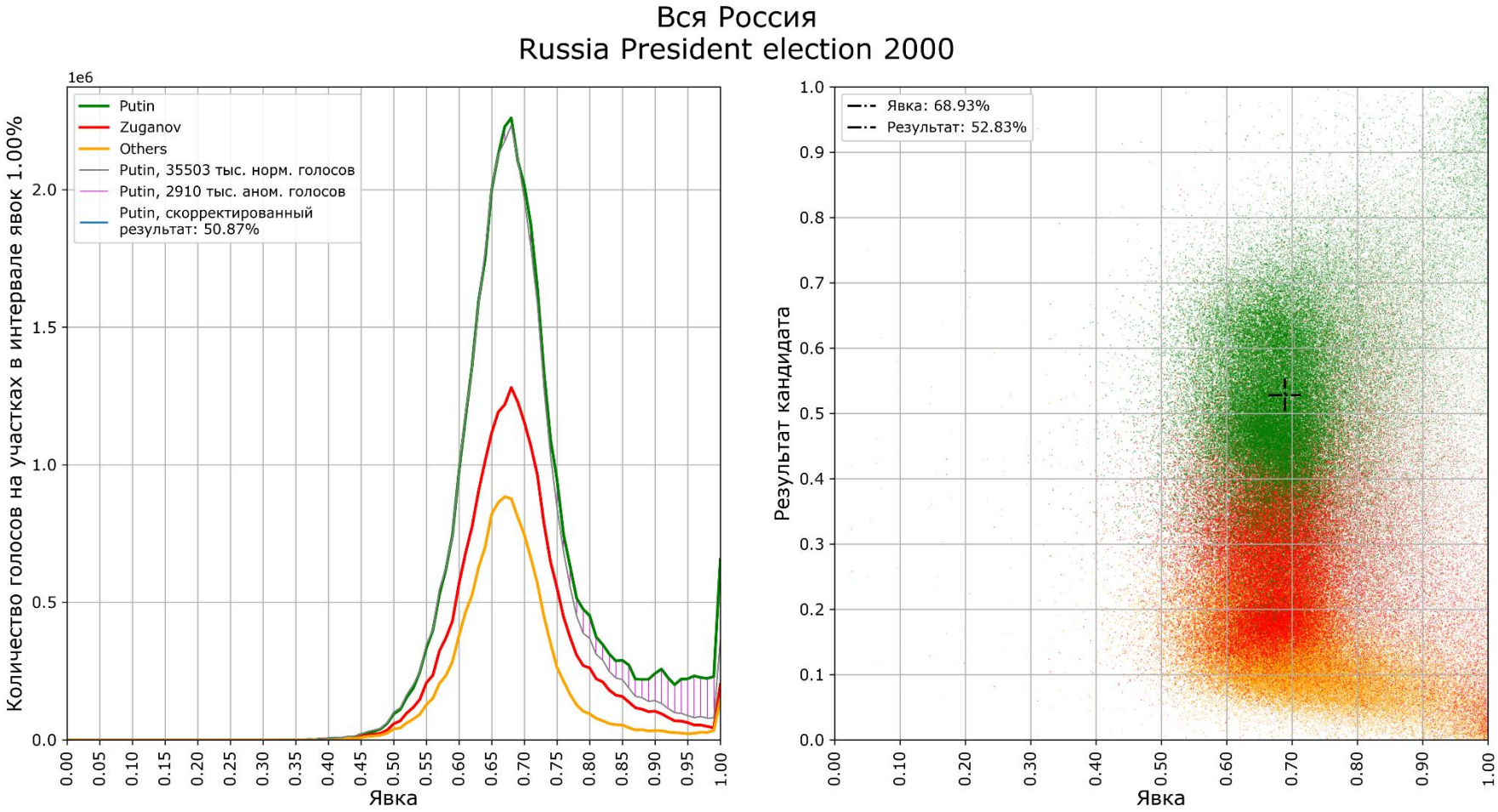
1996-1: -3.5 млн голосов



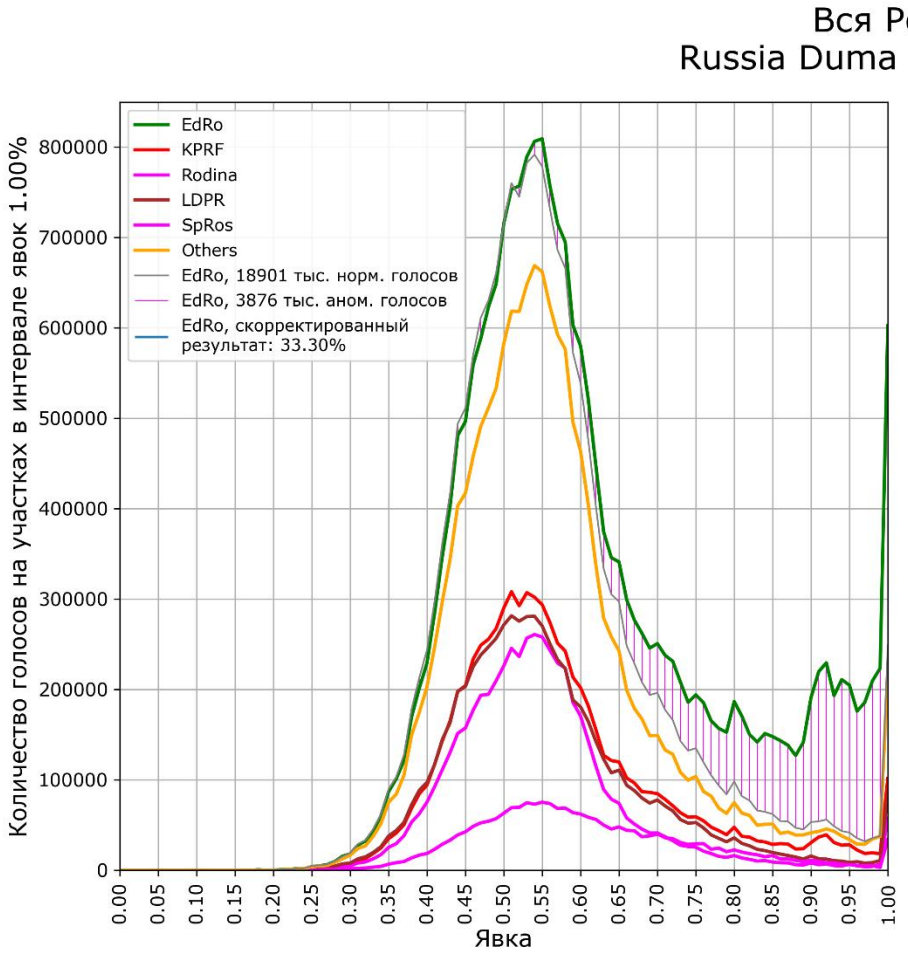
1996-2: -3.5 млн голосов

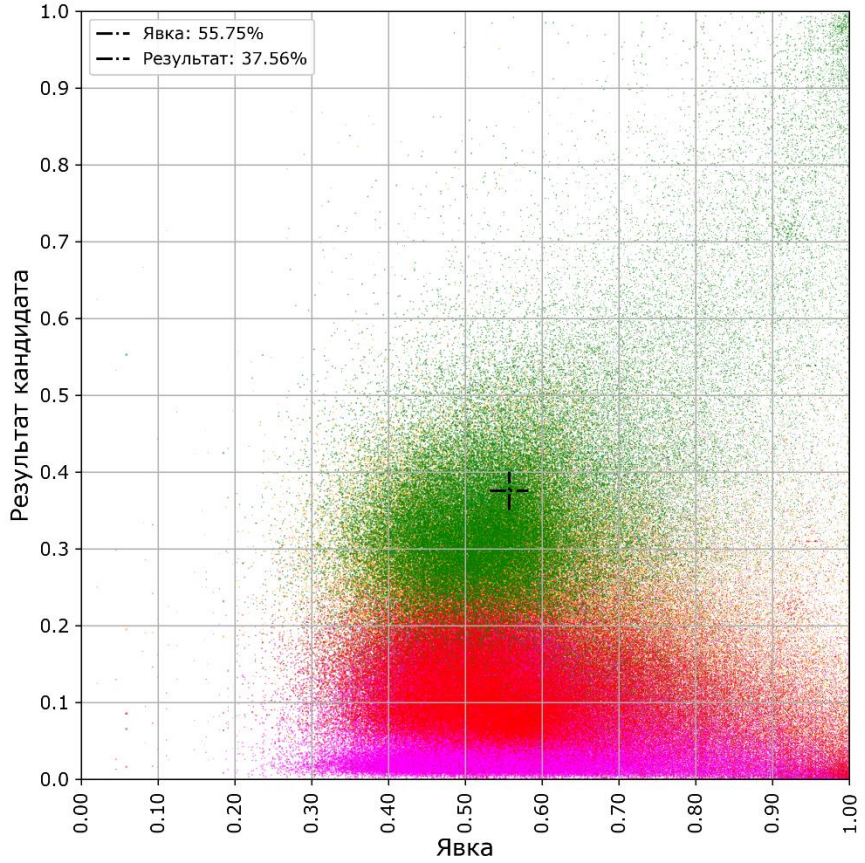


2000: 2.9 млн голосов

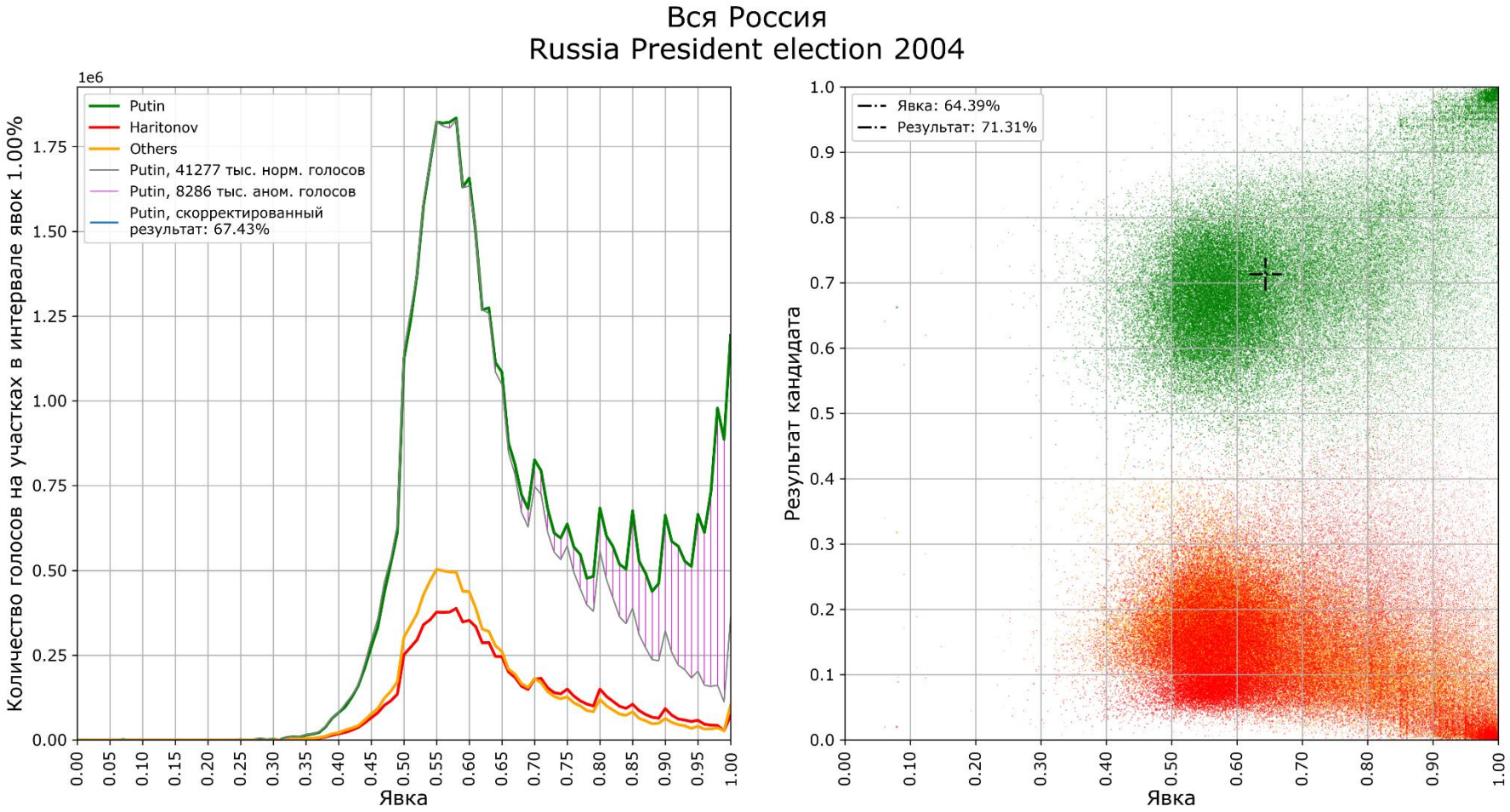


2003: 3.8 млн голосов

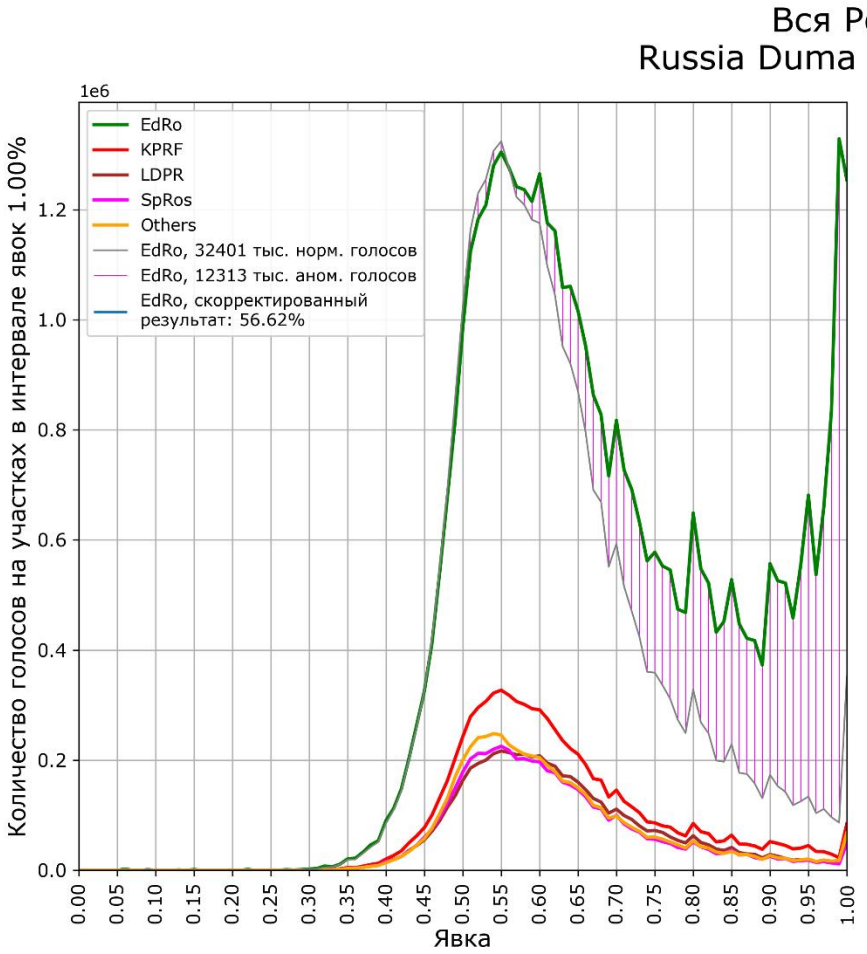


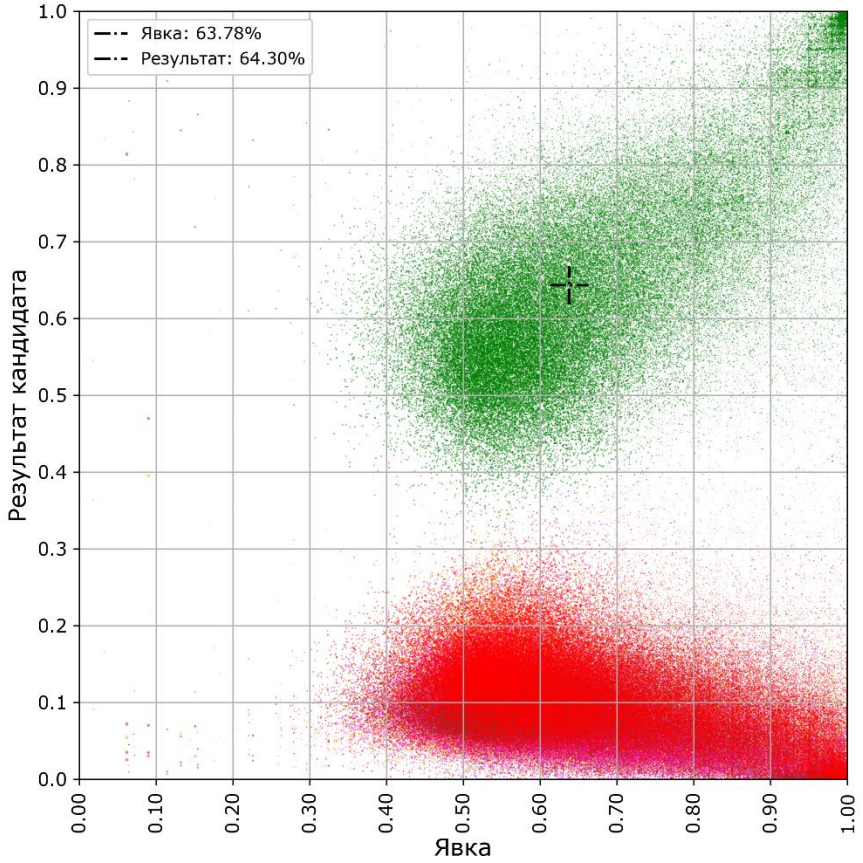


2004: 8.3 млн голосов

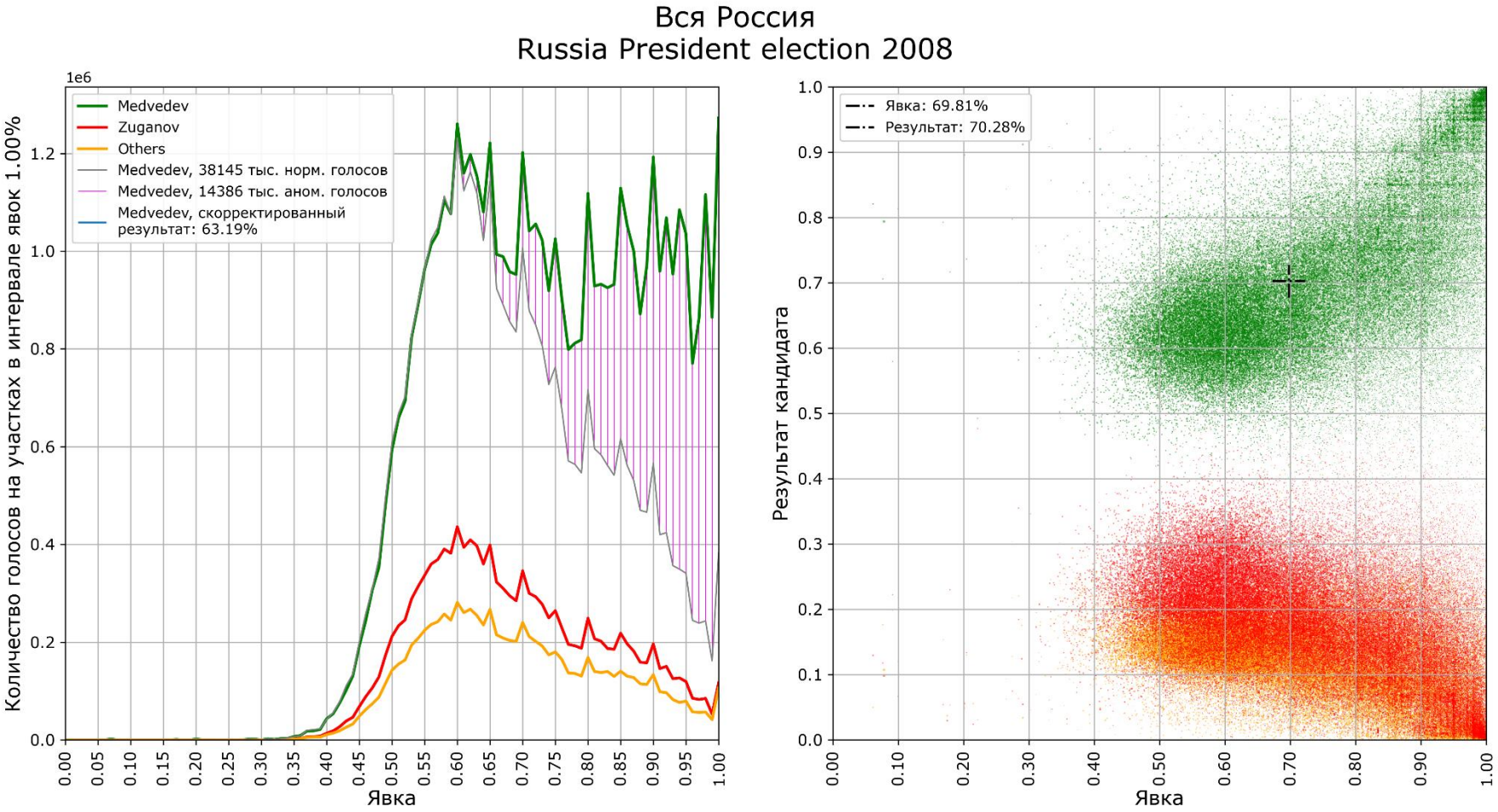


2007: 12.3 млн голосов

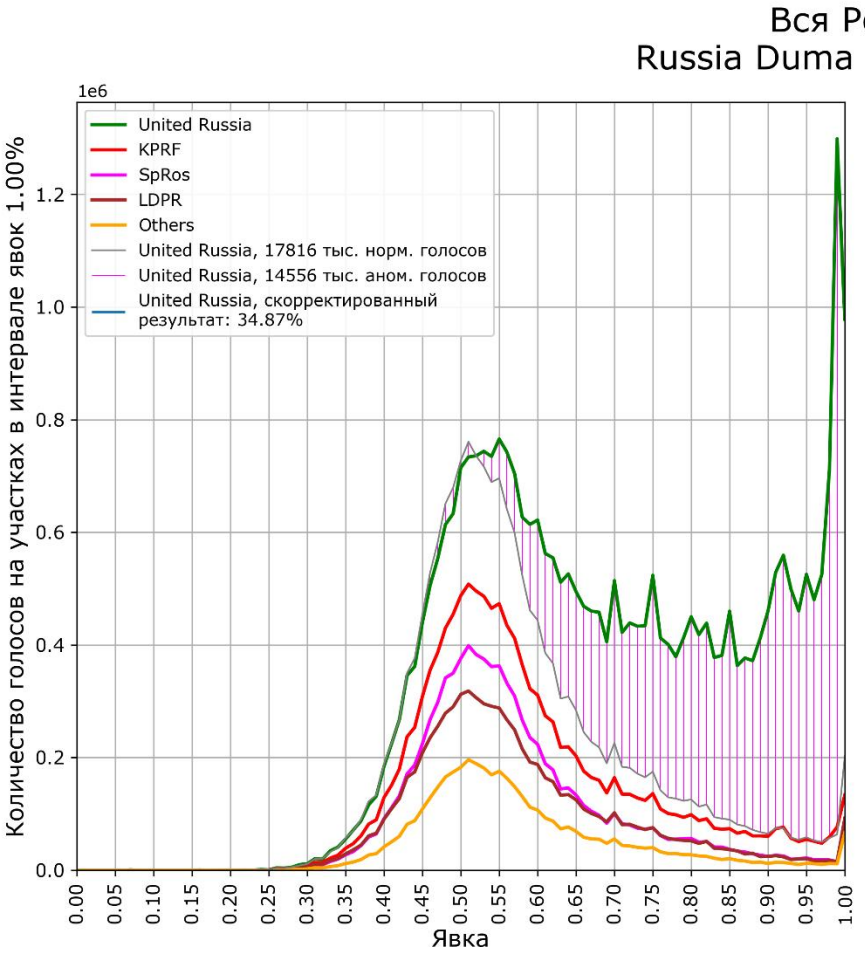


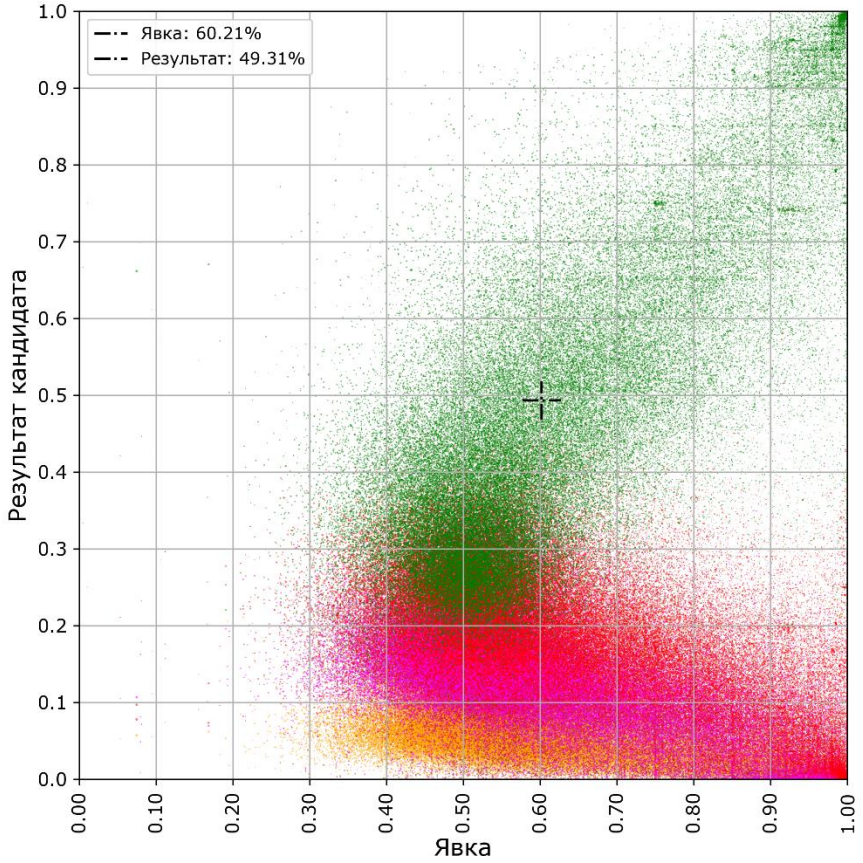


2008: 14.4 млн голосов

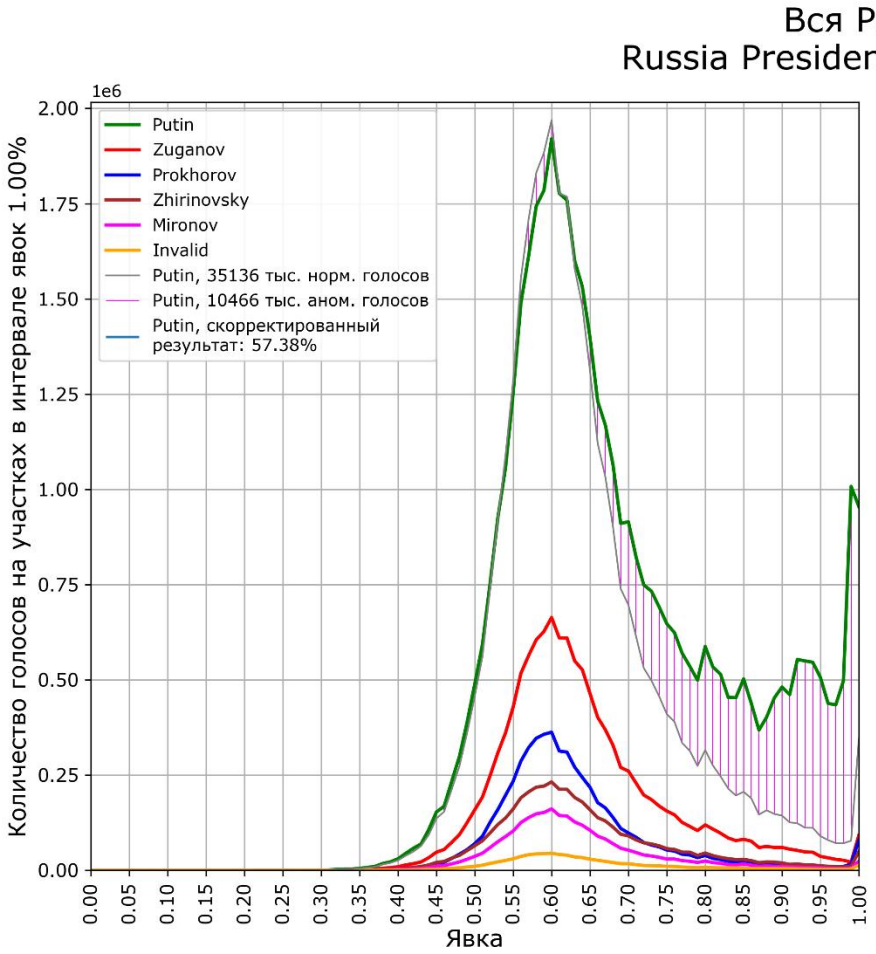


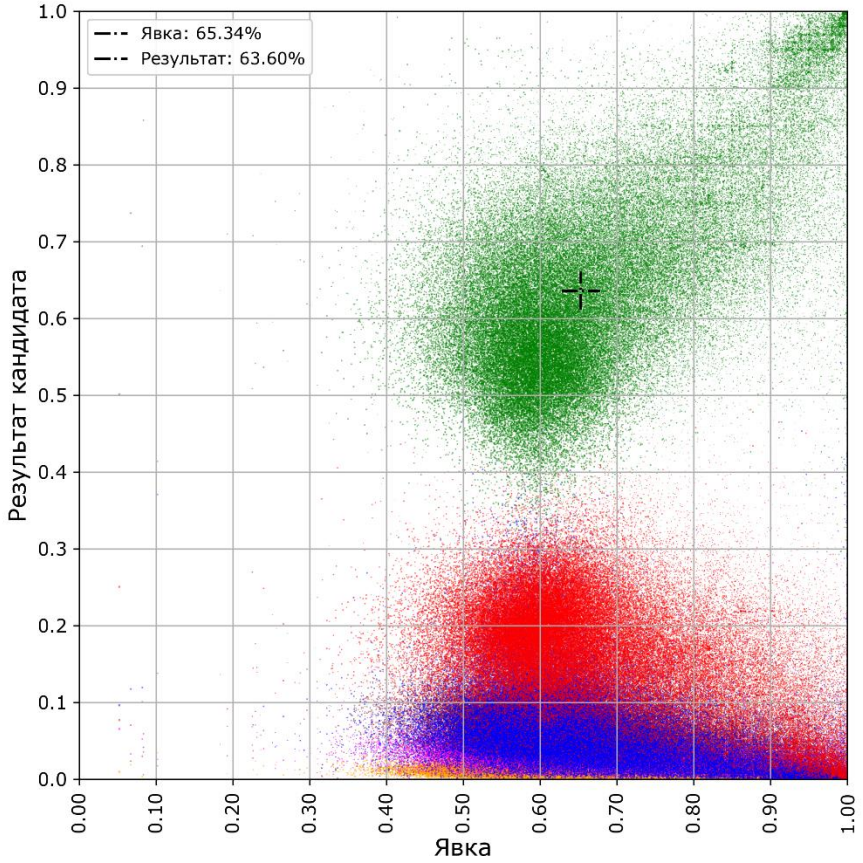
2011: 14.6 млн голосов



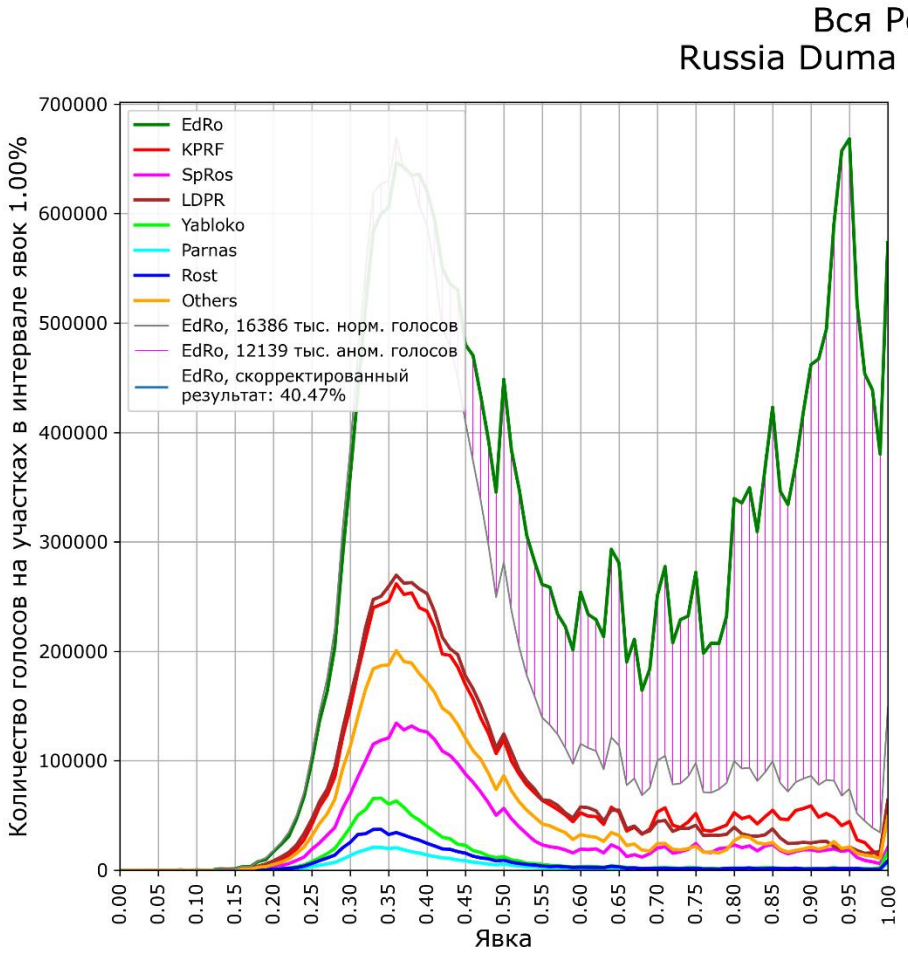


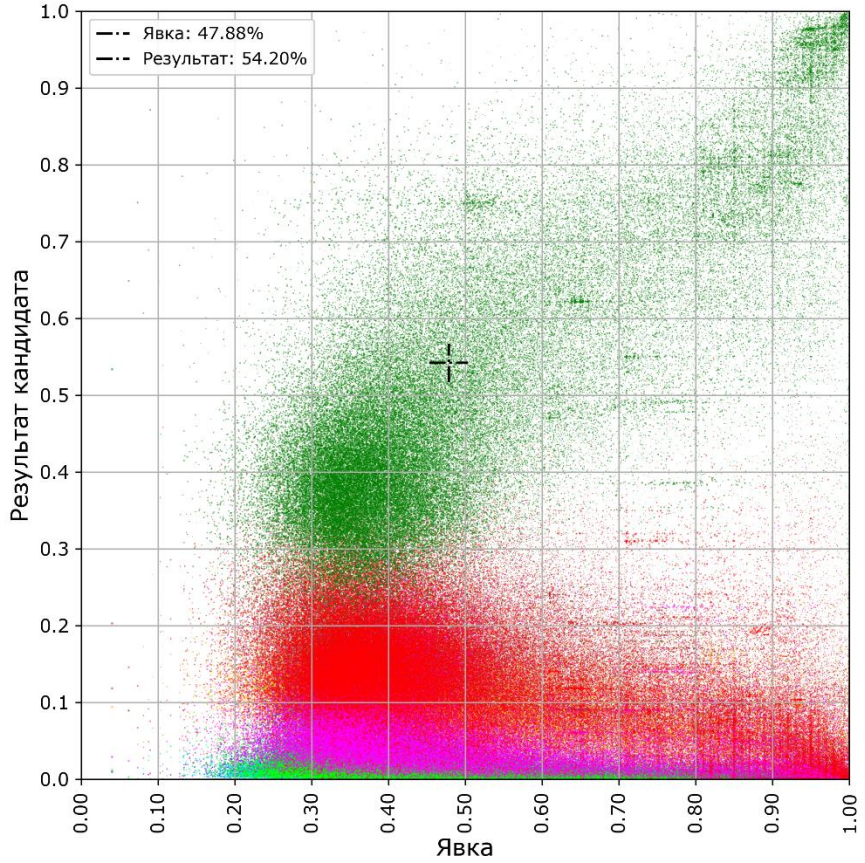
2012: 10.4 млн голосов



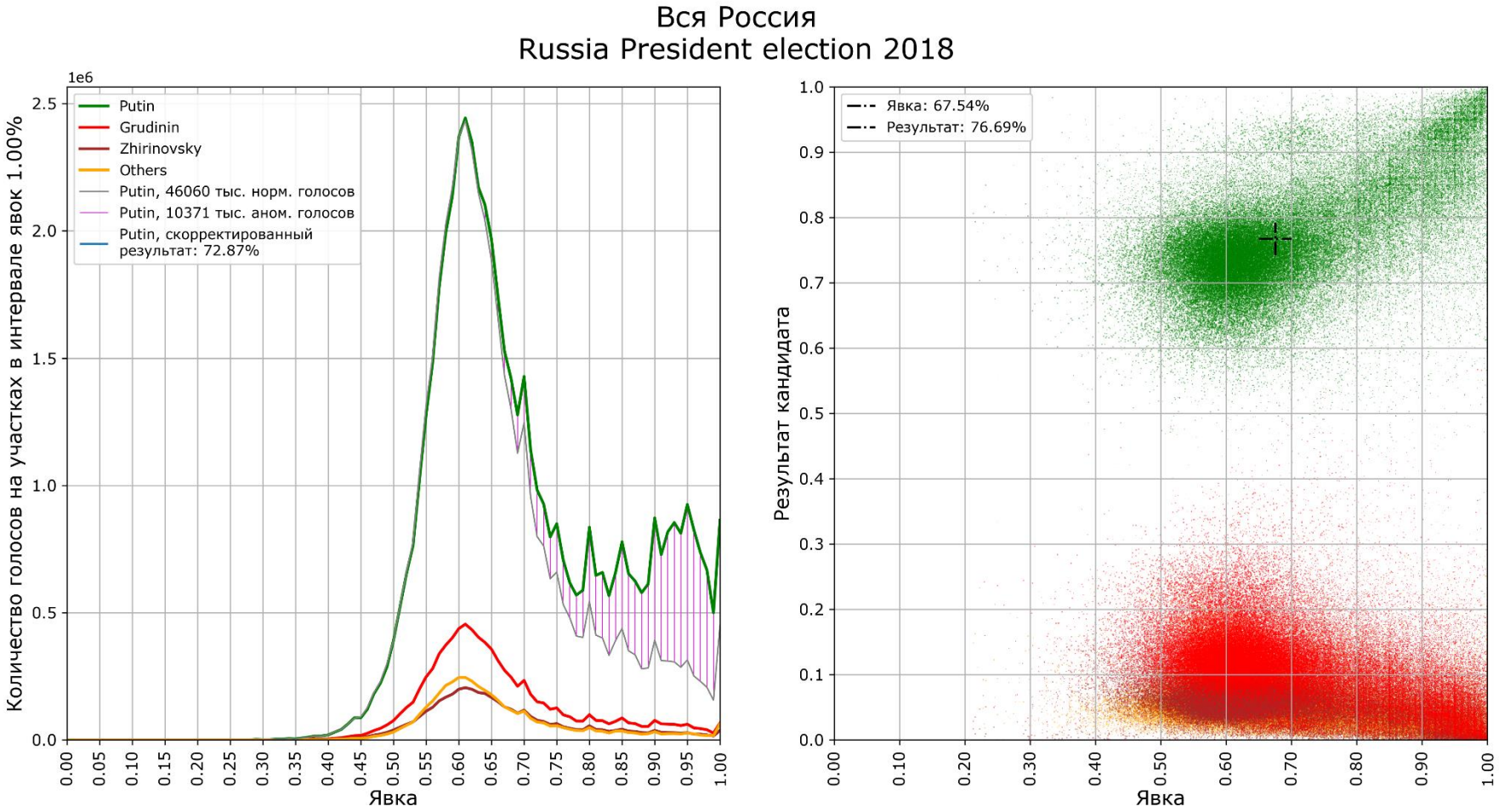


2016: 12.1 млн голосов

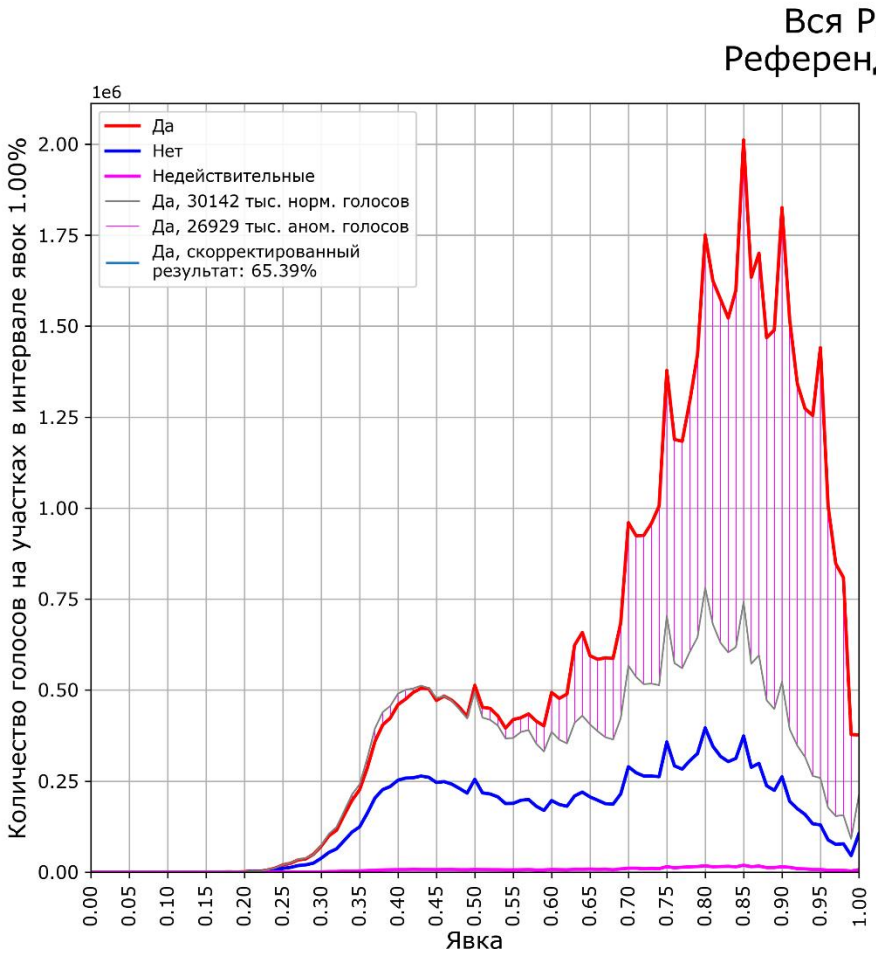


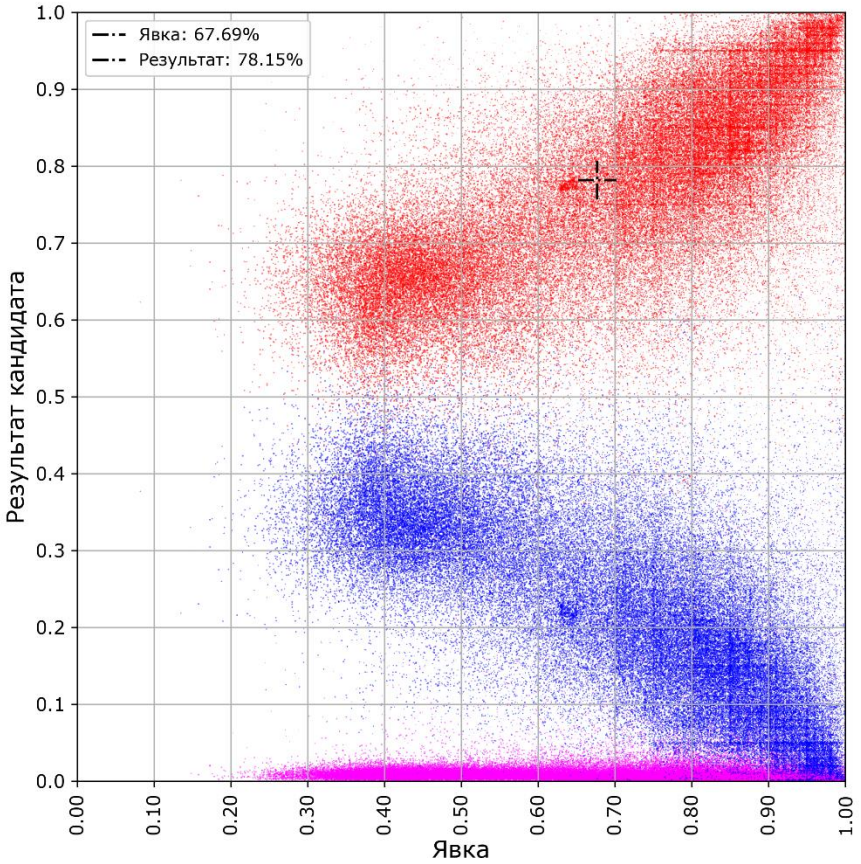


2018: 10.4 млн голосов

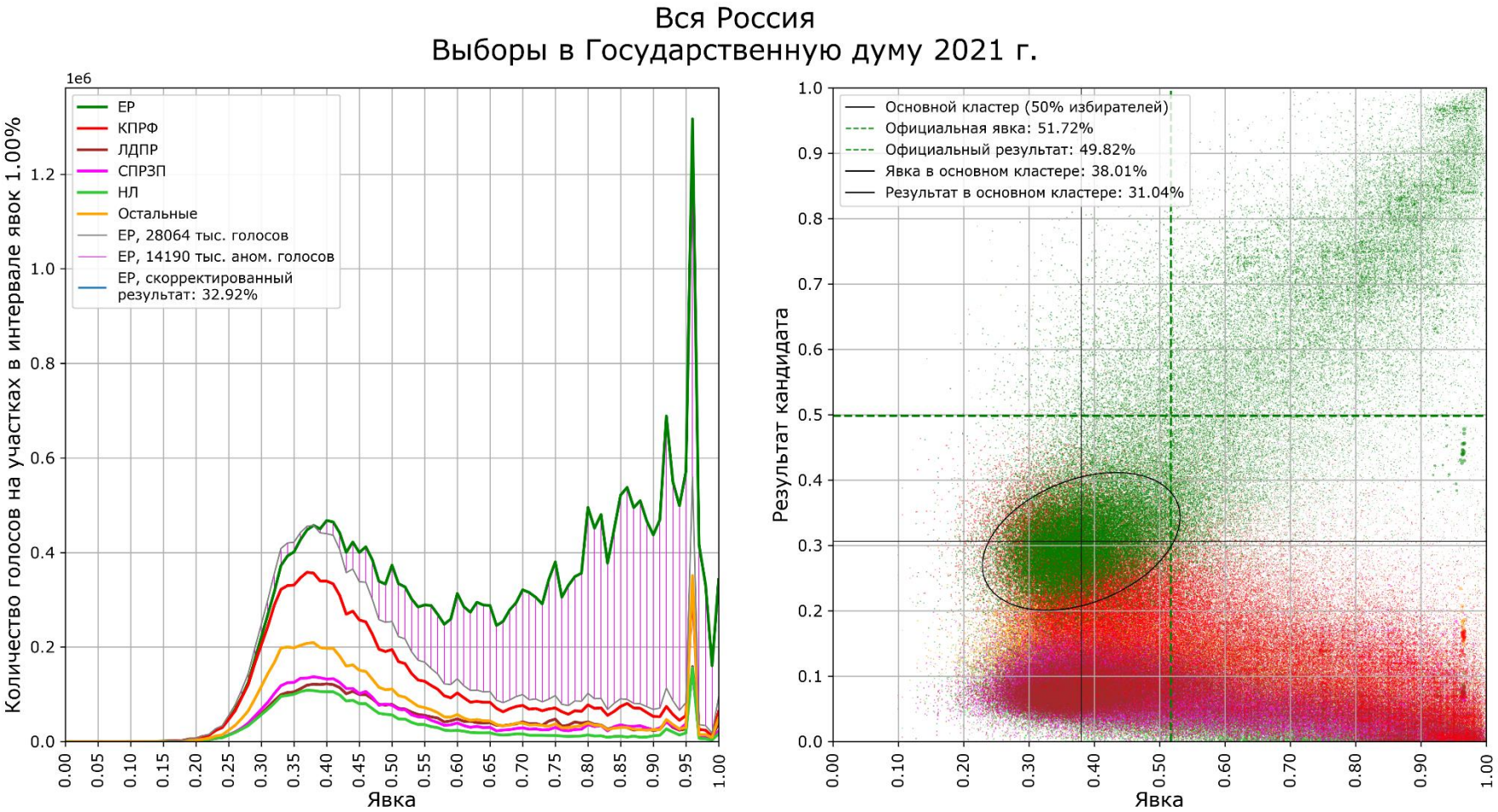


2020: 27 млн голосов

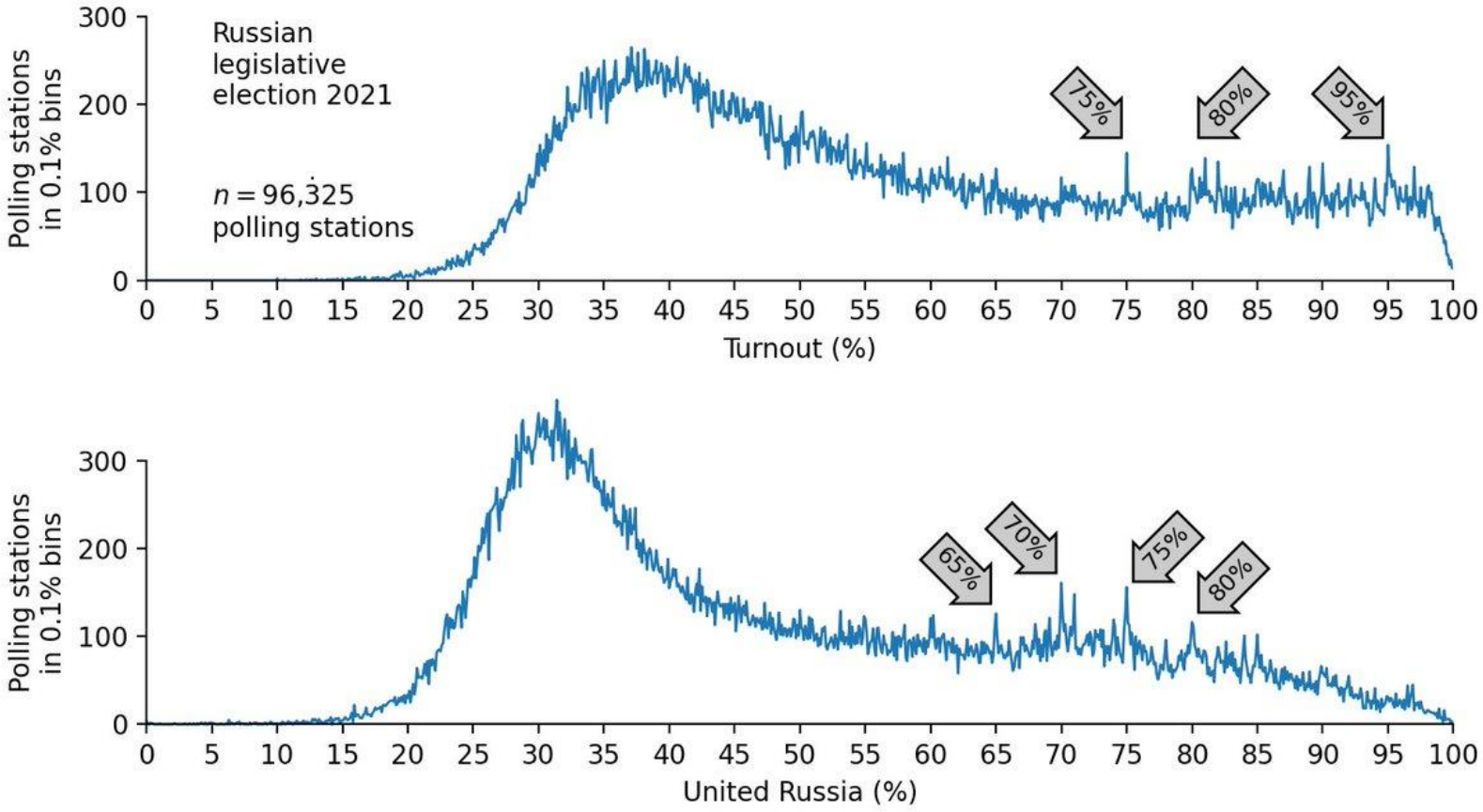




2021: 14.1 млн голосов

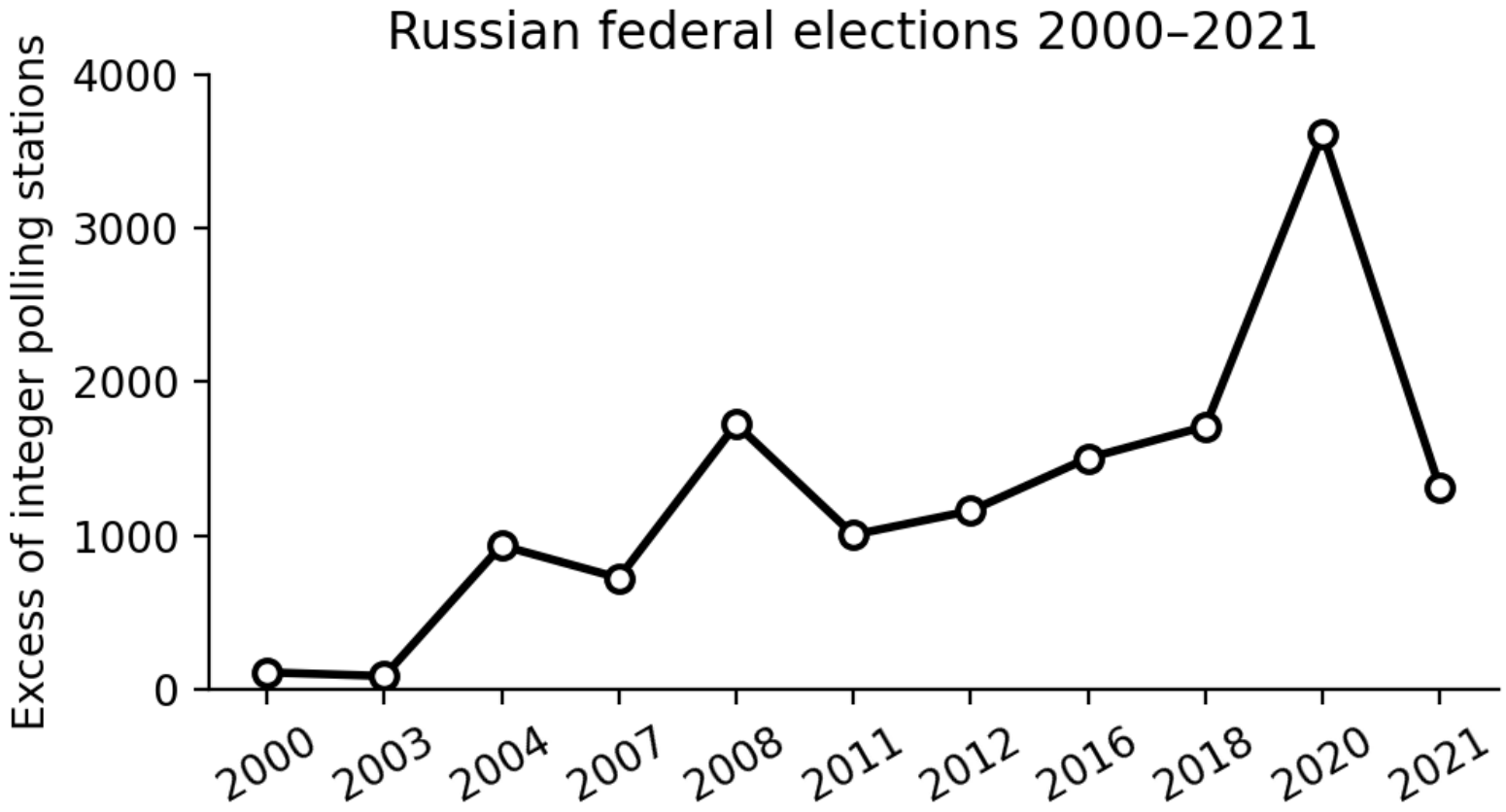


Пила Чурова



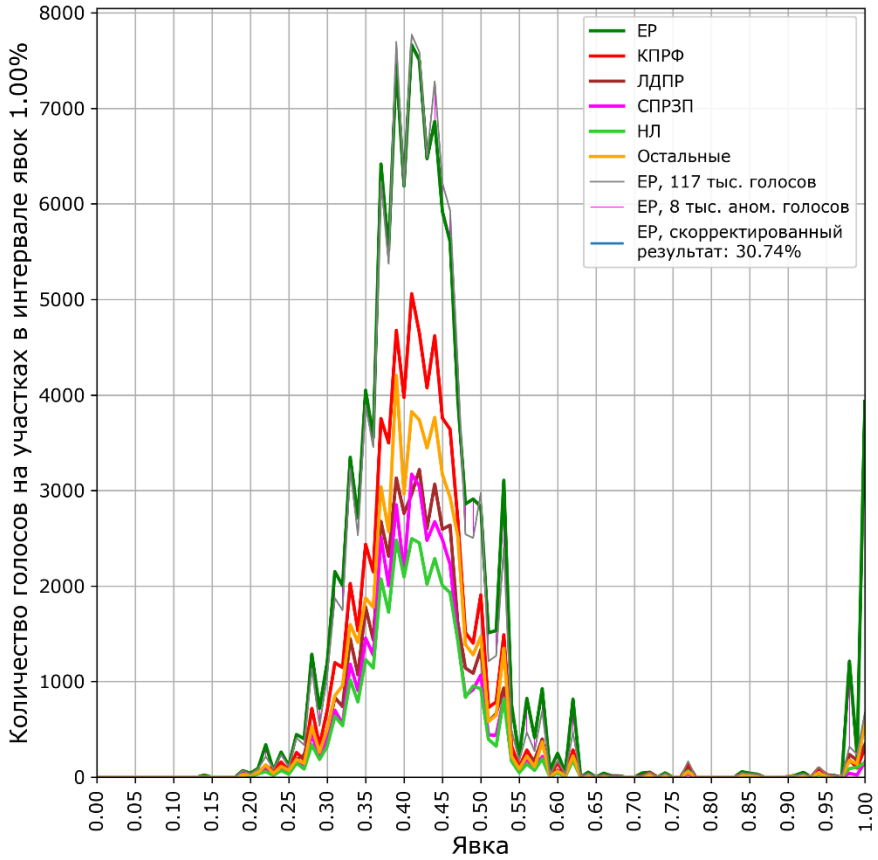
(Д. Кобак https://twitter.com/hippopedoid, 2021)

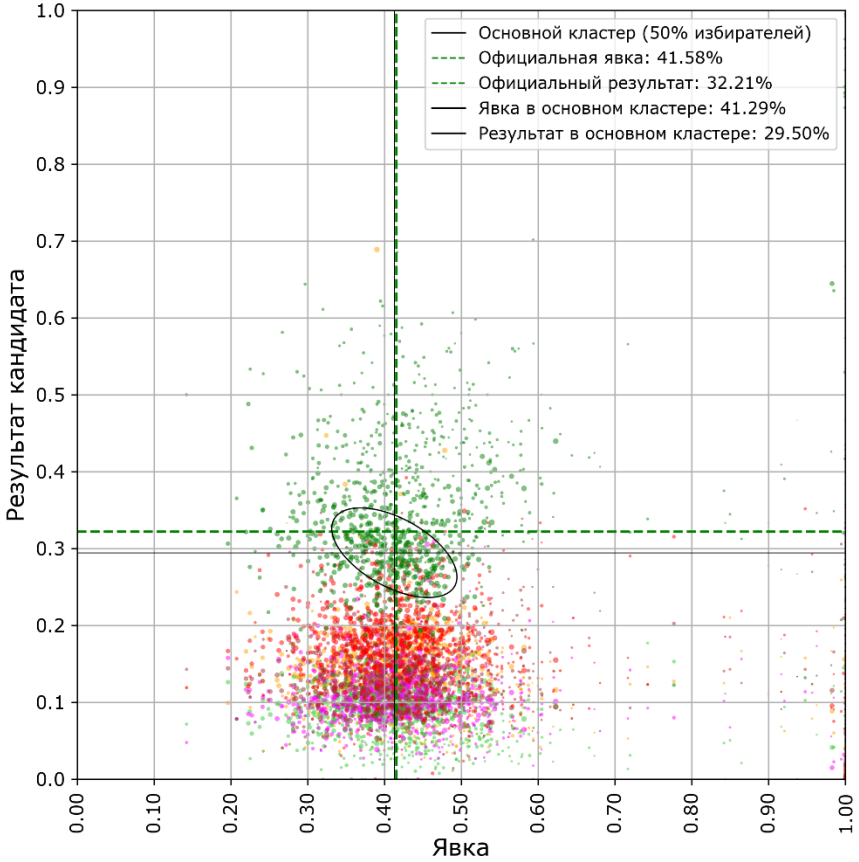
Пила Чурова



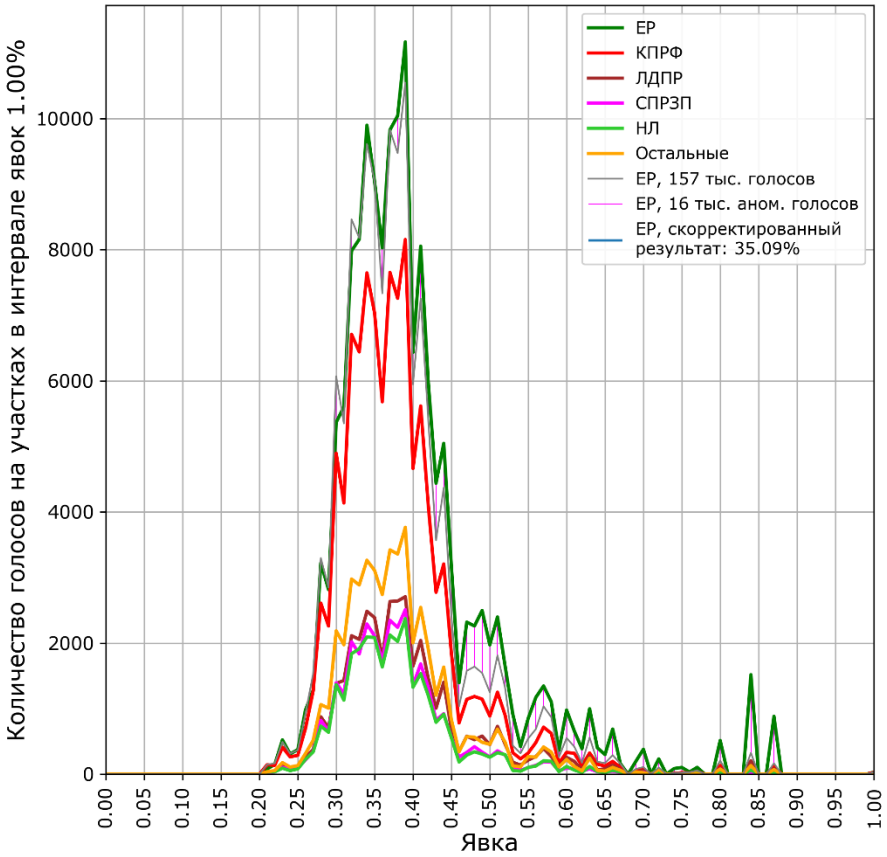
(Д. Кобак https://twitter.com/hippopedoid, 2021)

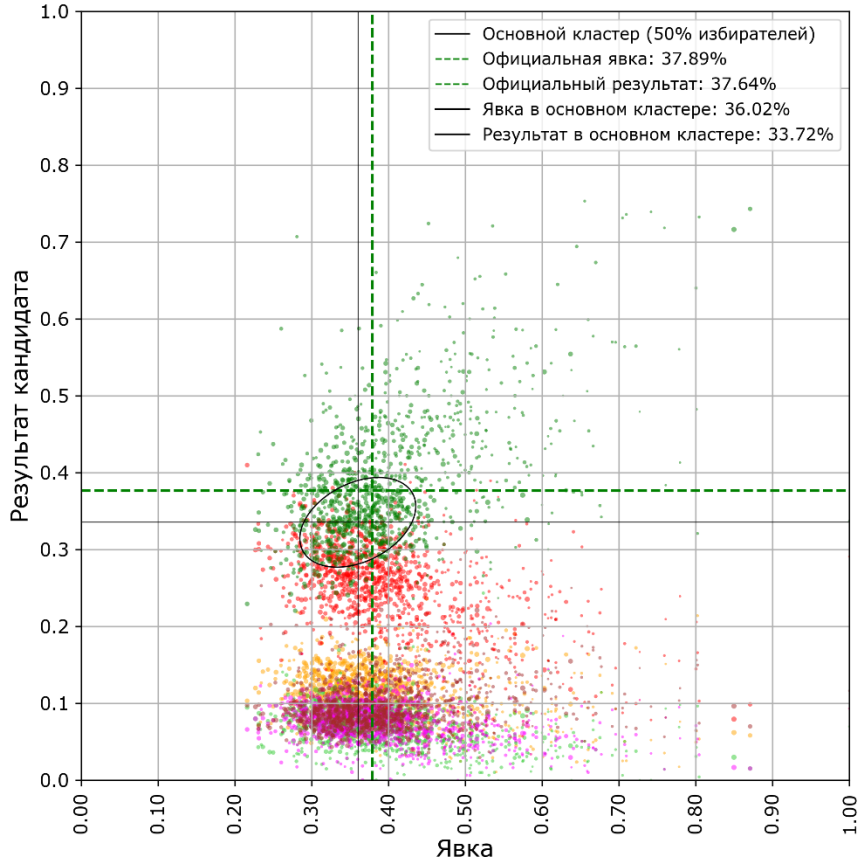
Архангельская область
Выборы в Государственную думу 2021 г.



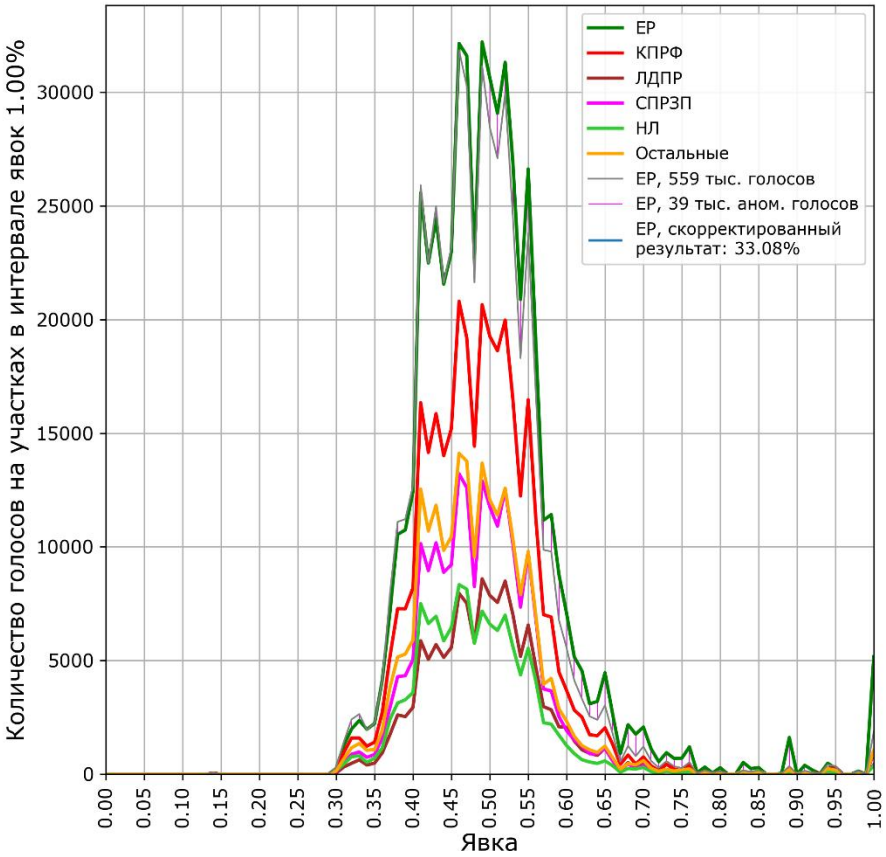


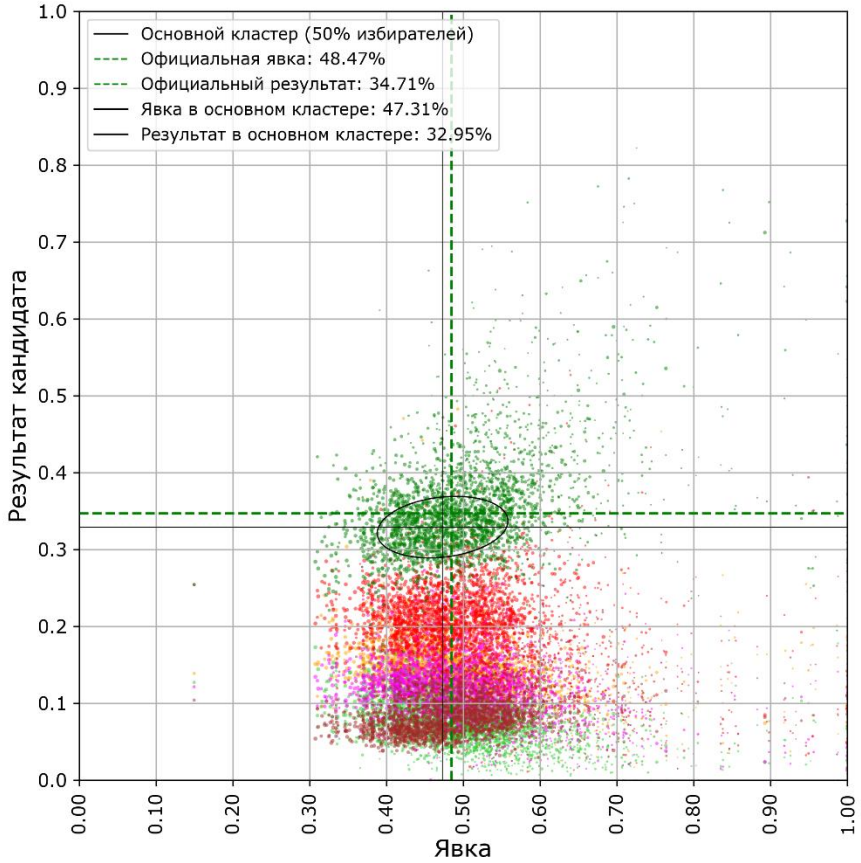
Владимирская область
Выборы в Государственную думу 2021 г.



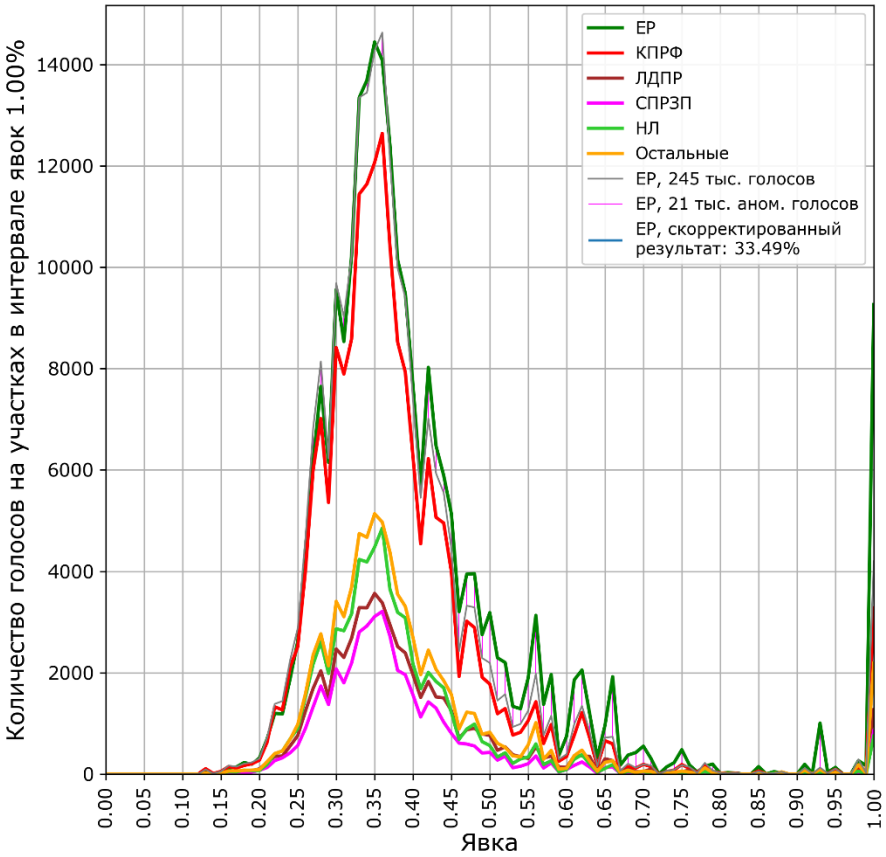


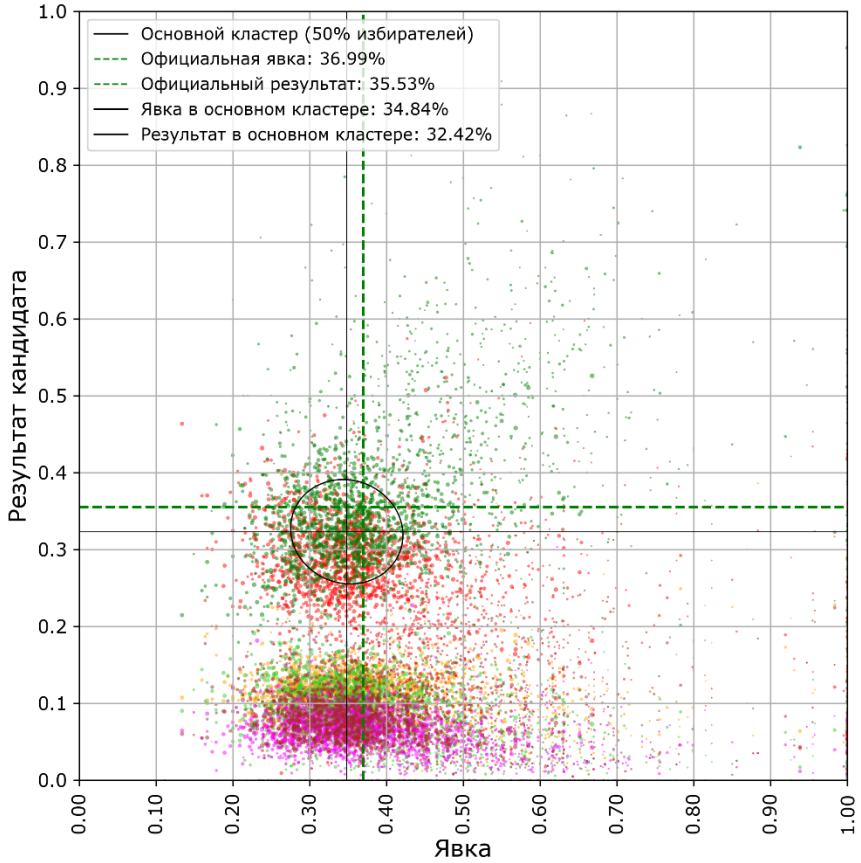
Свердловская область
Выборы в Государственную думу 2021 г.



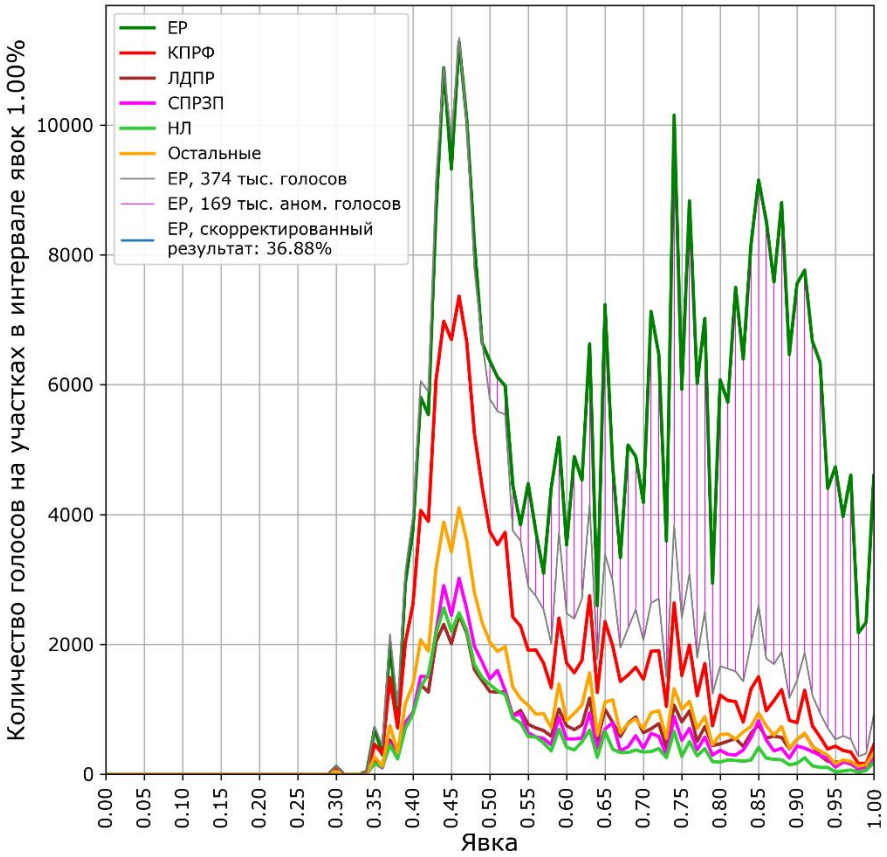


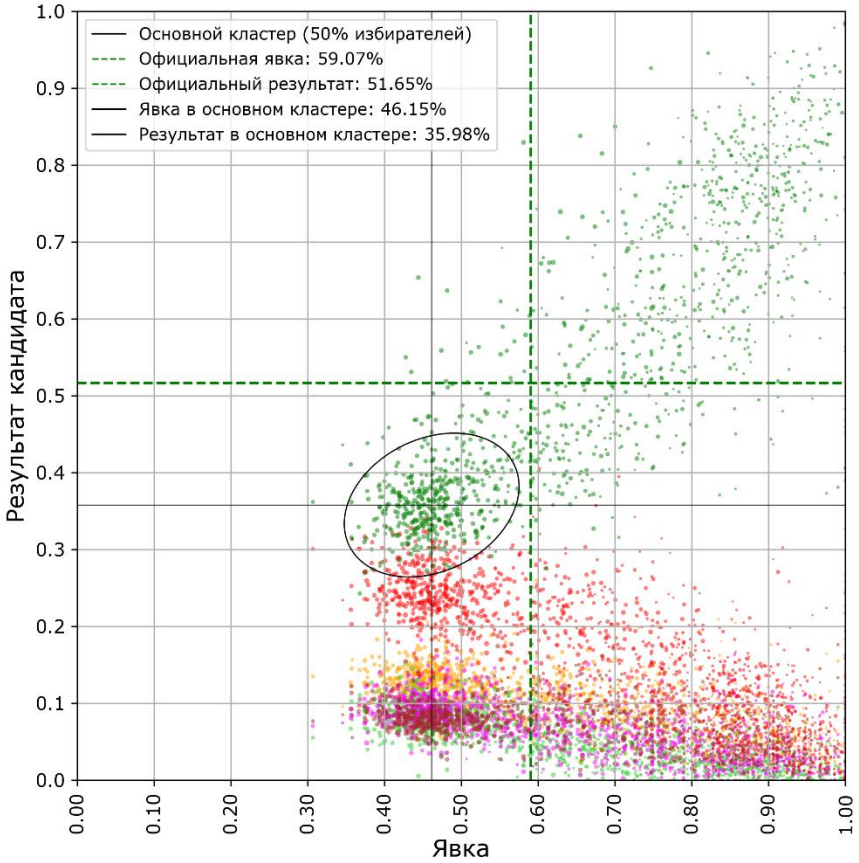
Иркутская область
Выборы в Государственную думу 2021 г.



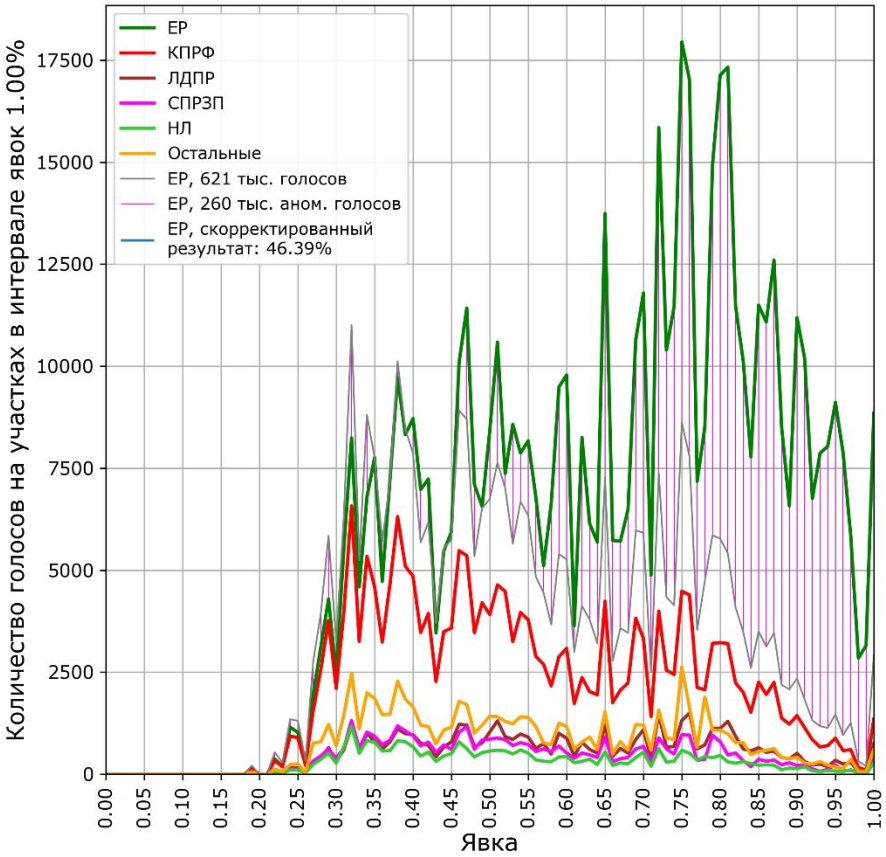


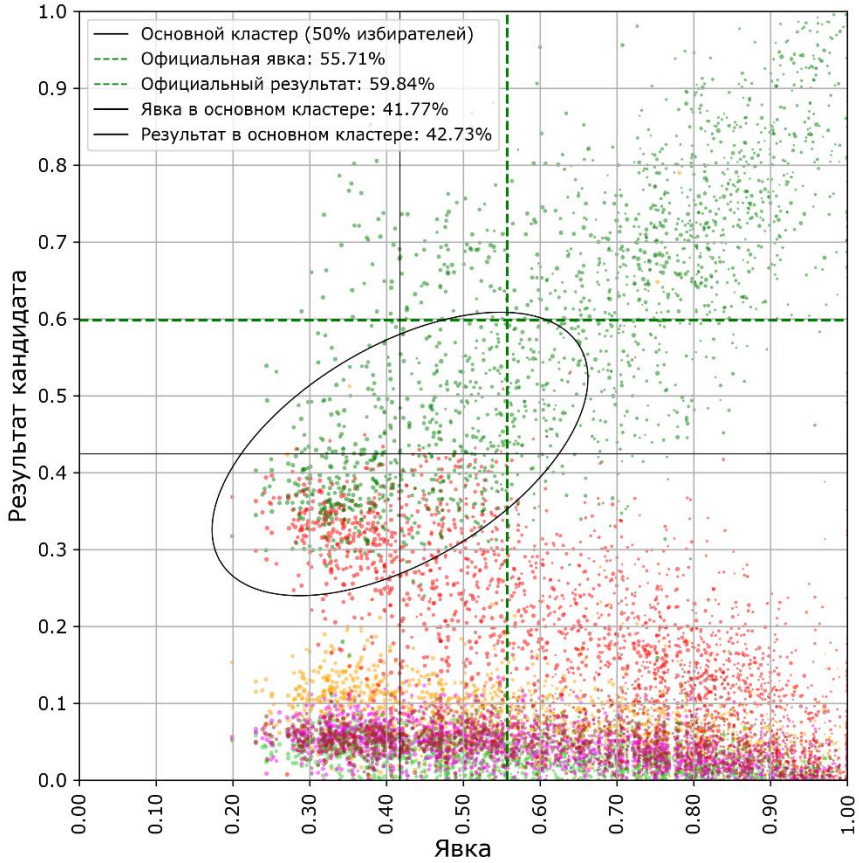
Белгородская область
Выборы в Государственную думу 2021 г.



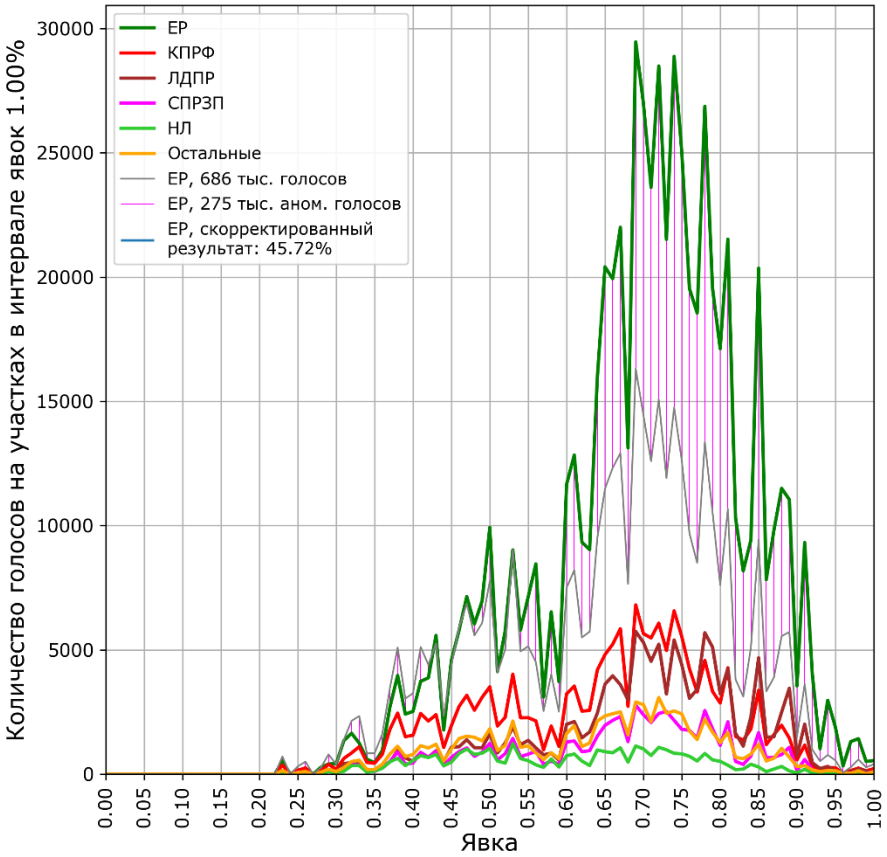


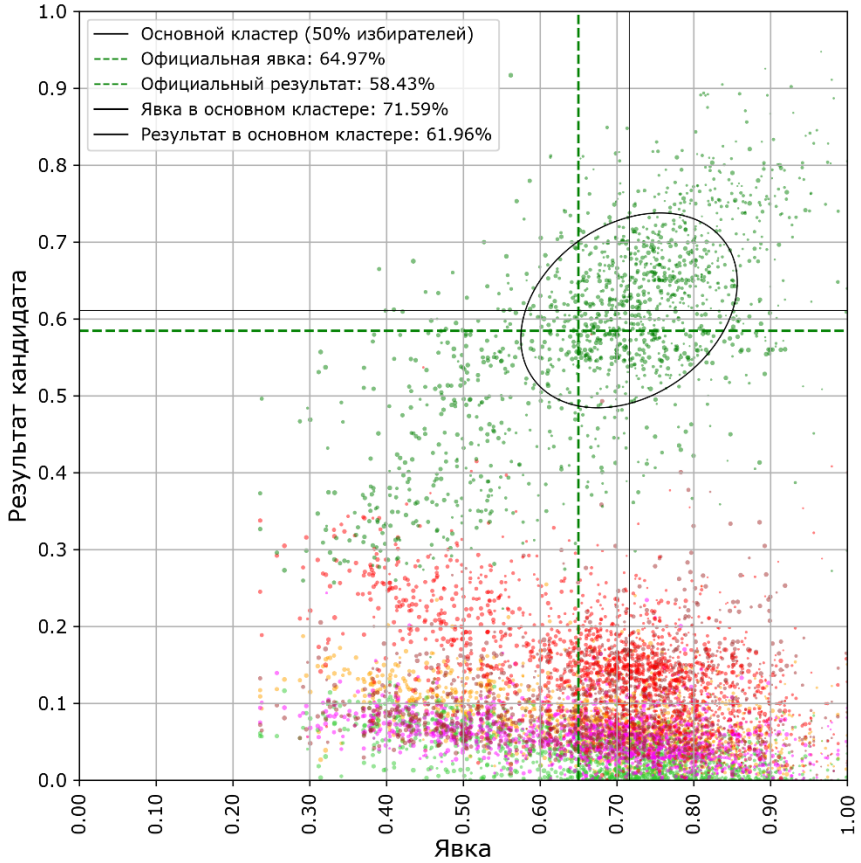
Саратовская область
Выборы в Государственную думу 2021 г.



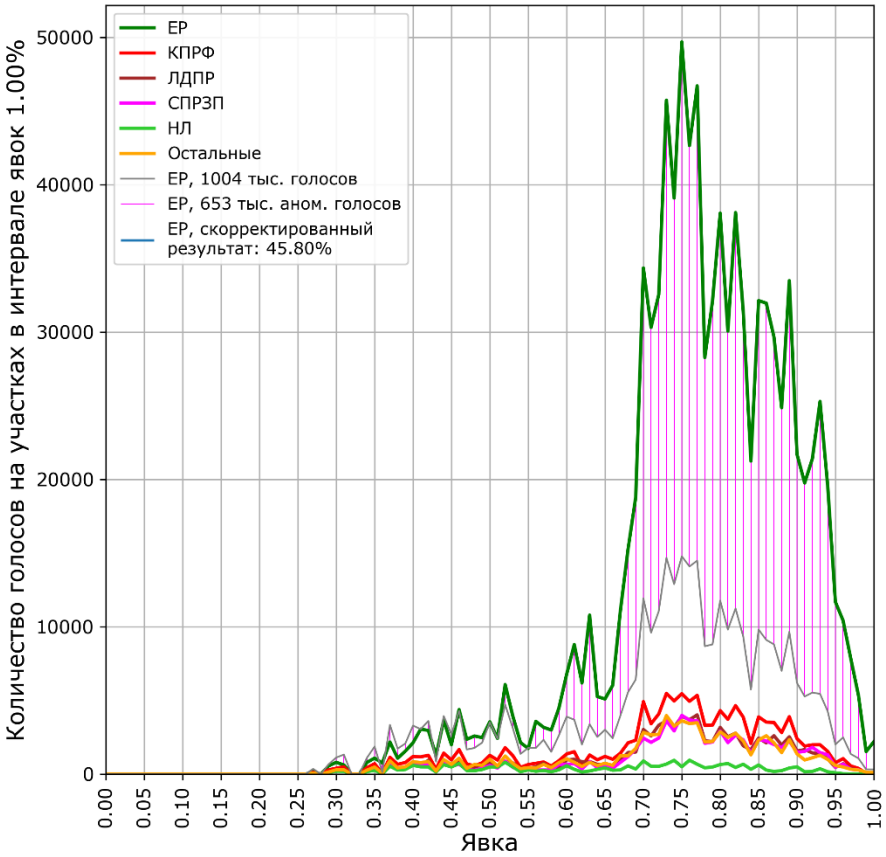


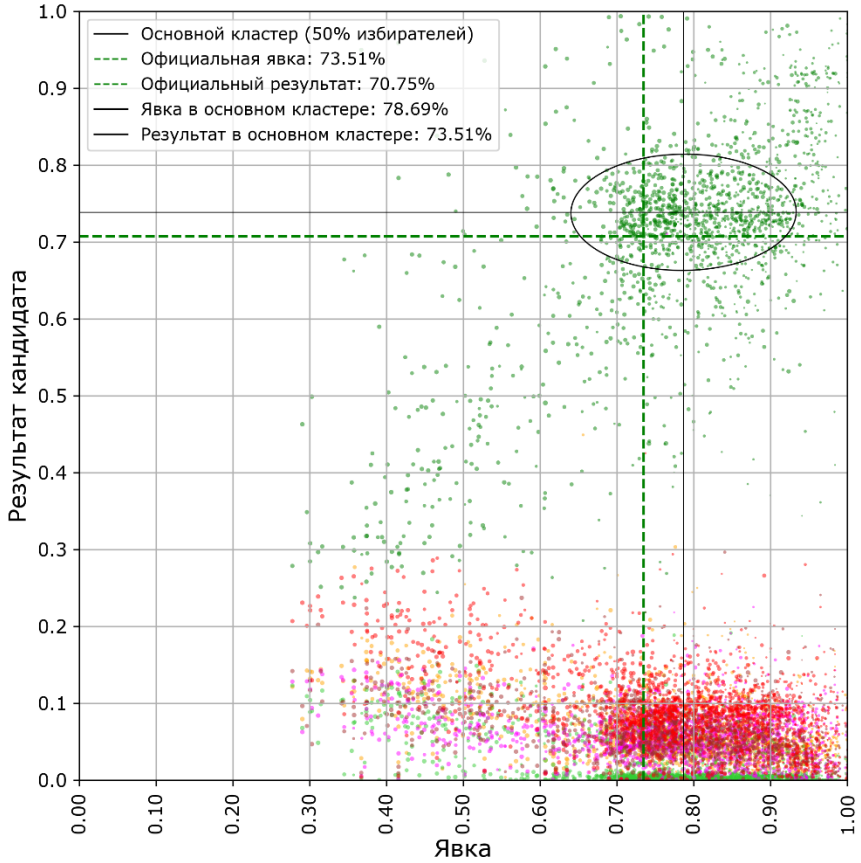
Волгоградская область
Выборы в Государственную думу 2021 г.



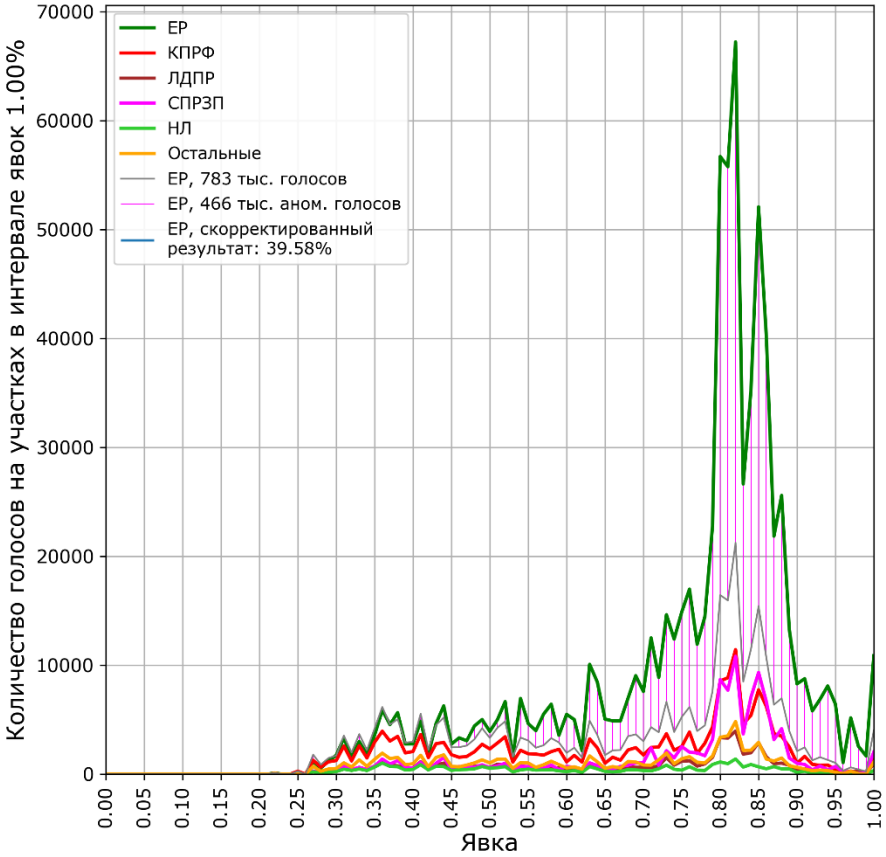


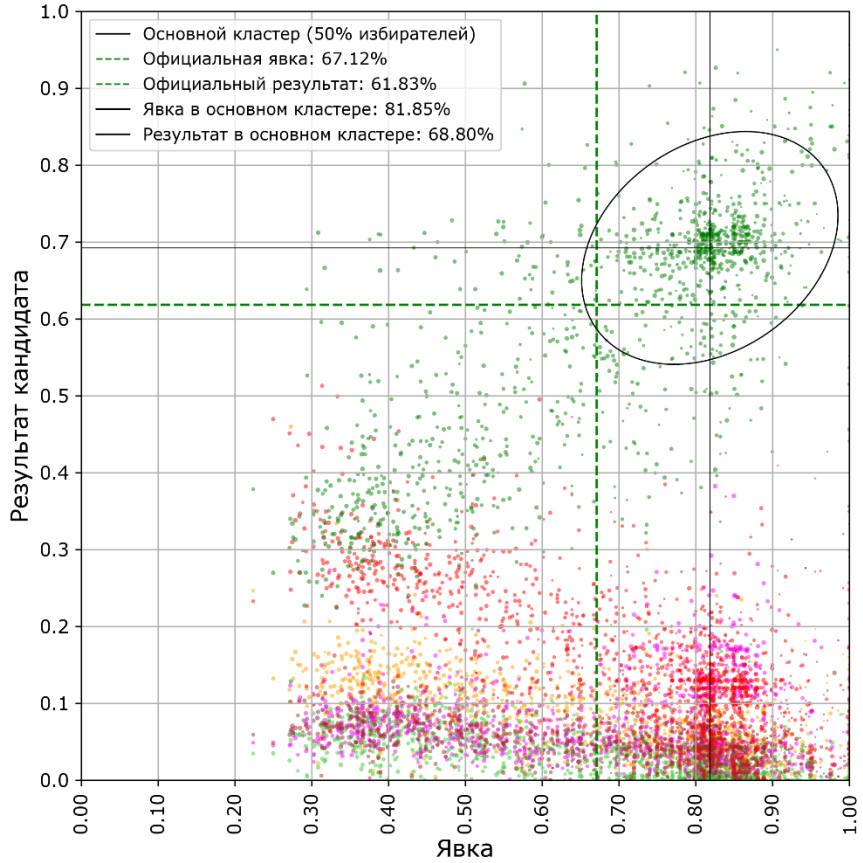
Кемеровская область - Кузбасс
Выборы в Государственную думу 2021 г.



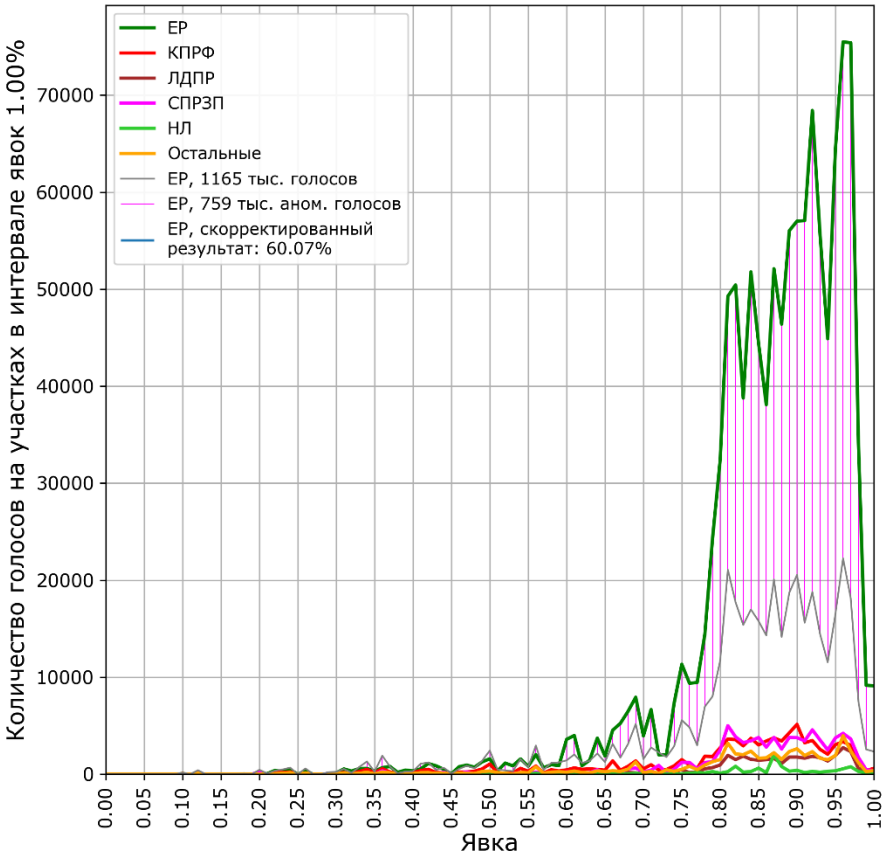


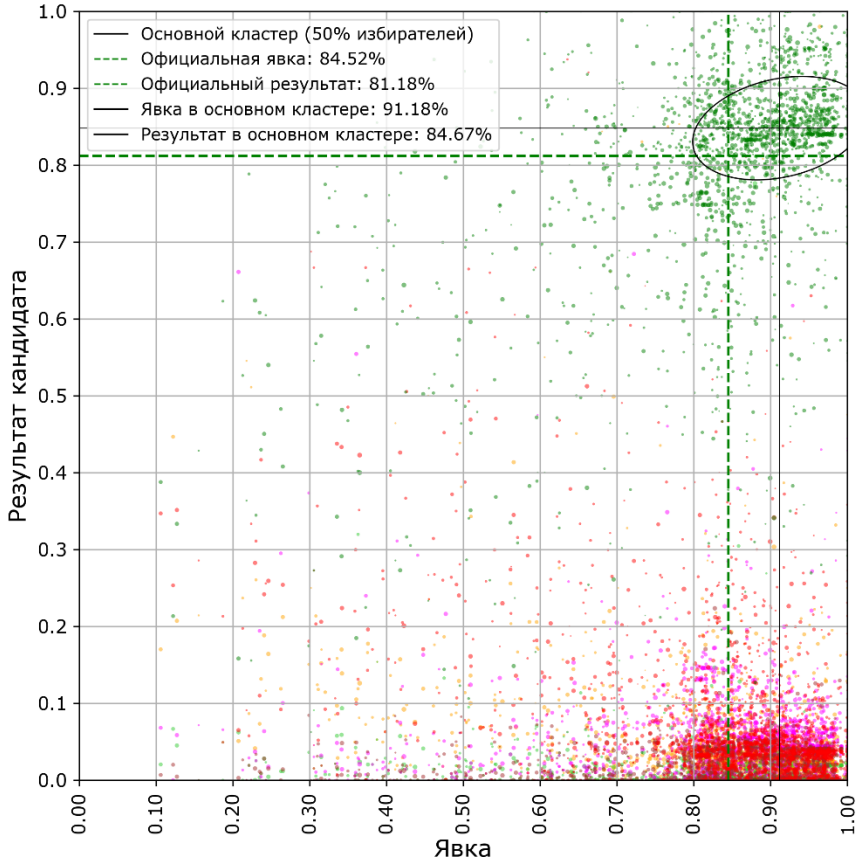
Ставропольский край
Выборы в Государственную думу 2021 г.



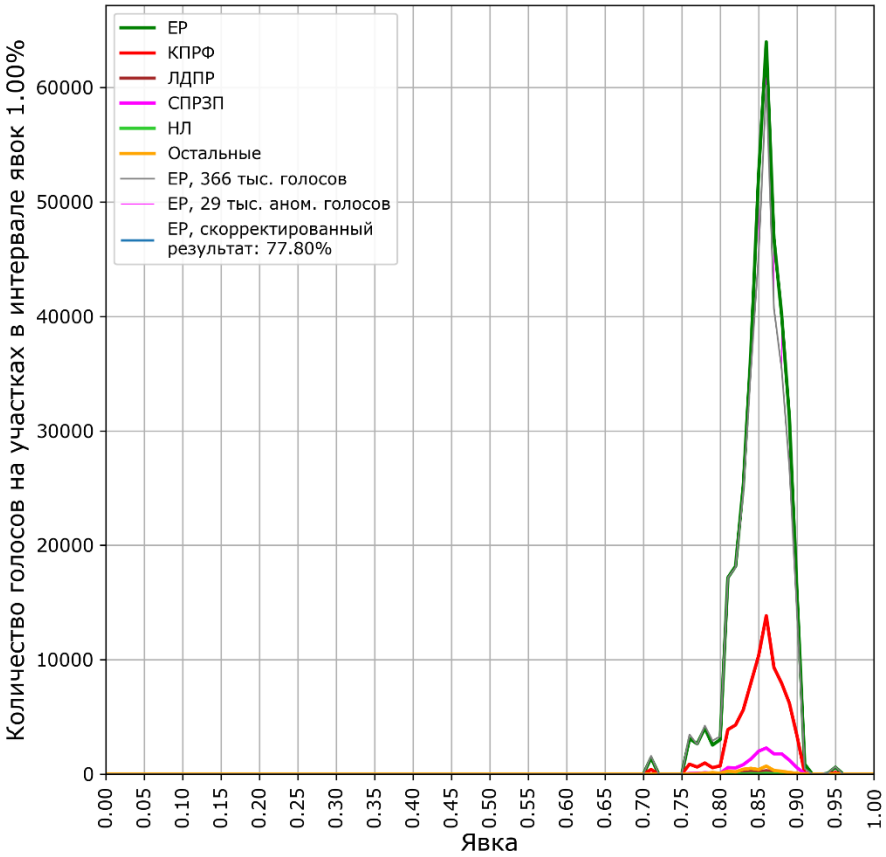


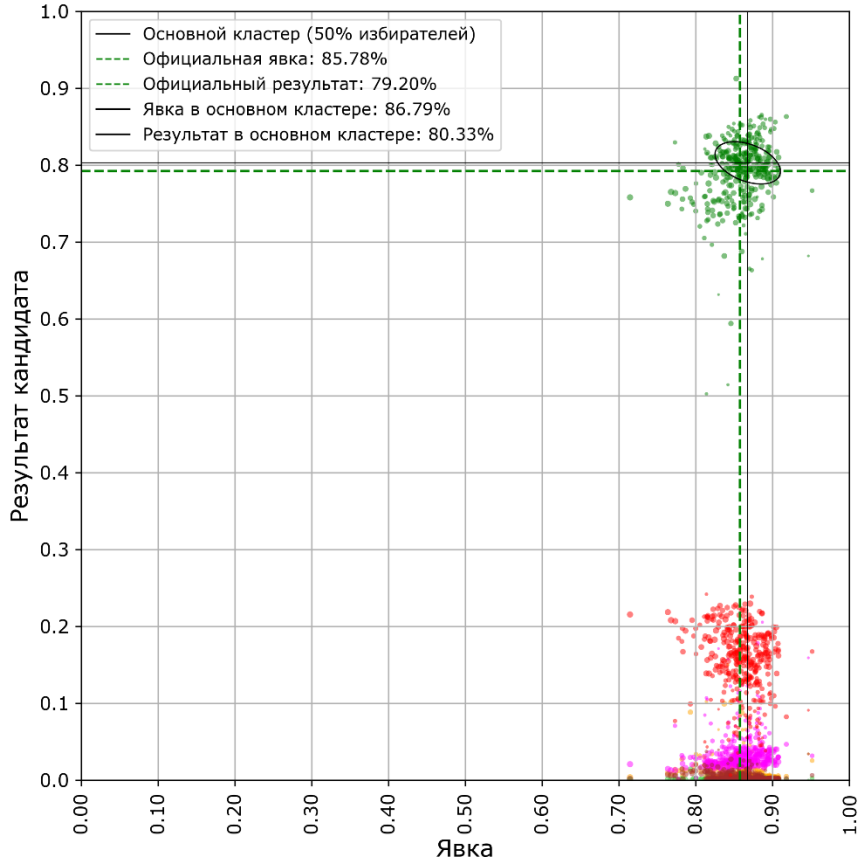
Республика Дагестан
Выборы в Государственную думу 2021 г.



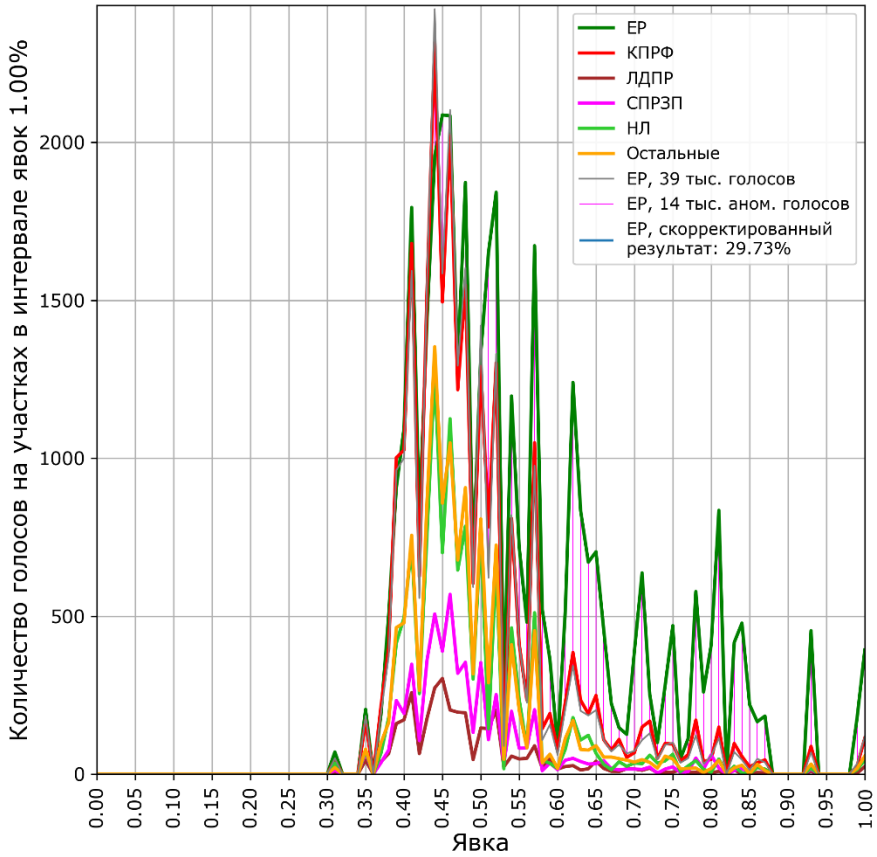


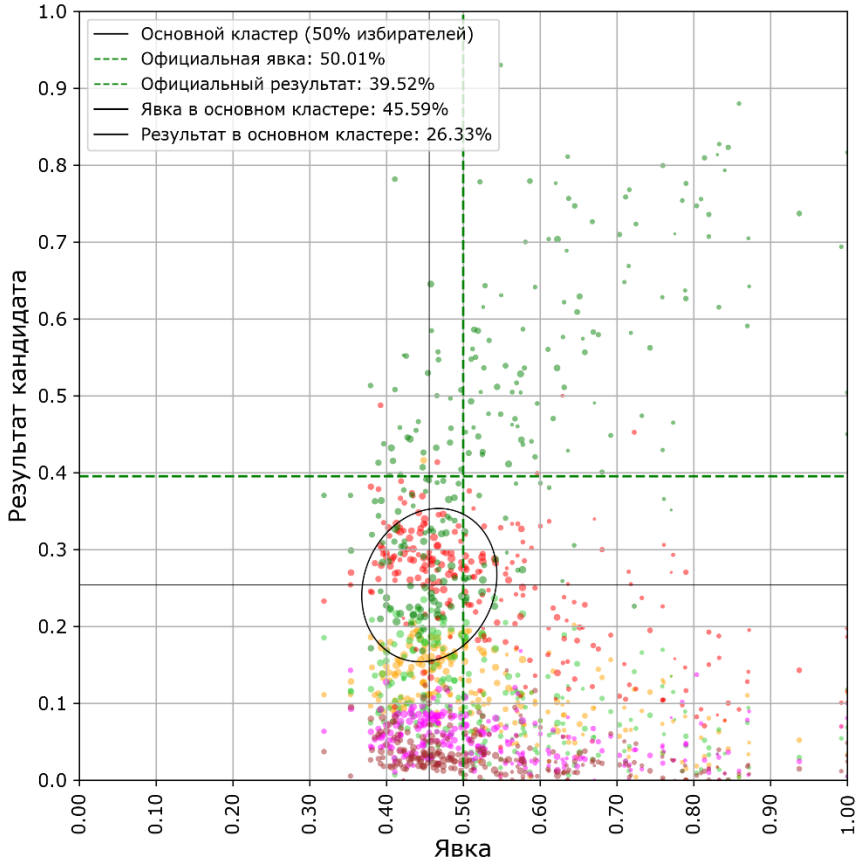
Кабардино-Балкарская Республика
Выборы в Государственную думу 2021 г.



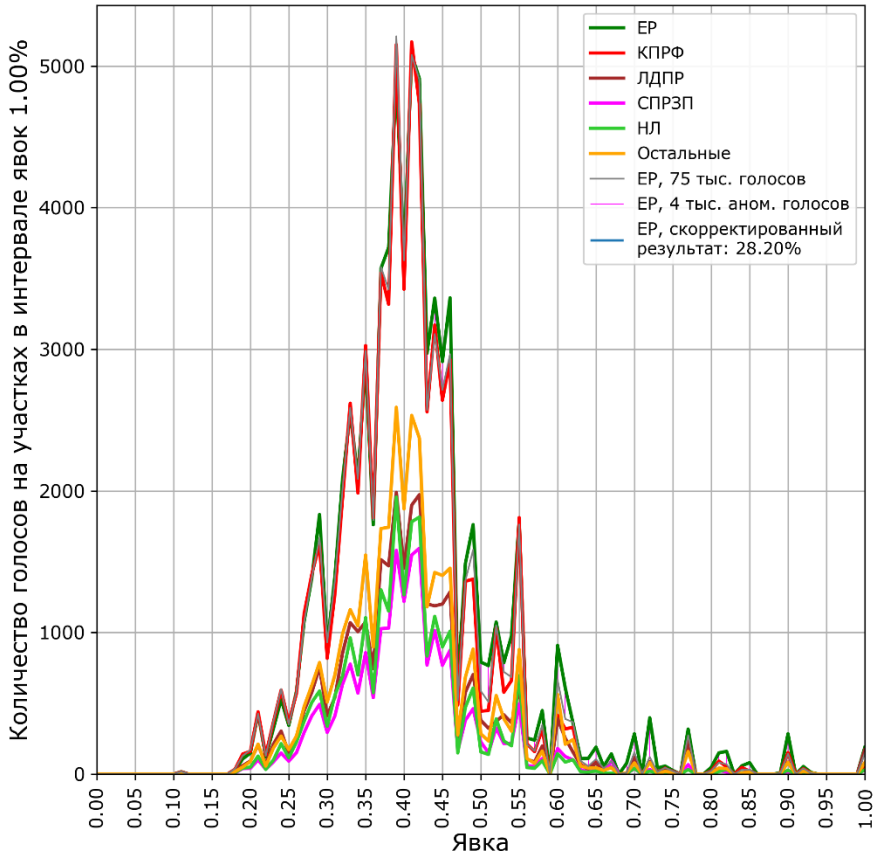


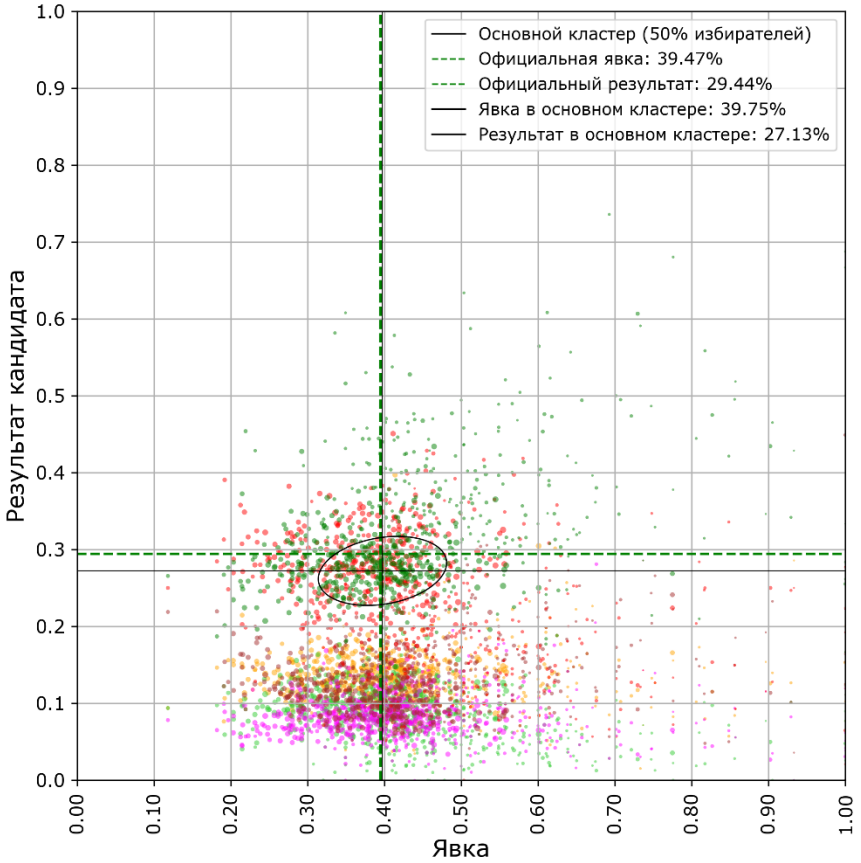
Республика Калмыкия
Выборы в Государственную думу 2021 г.



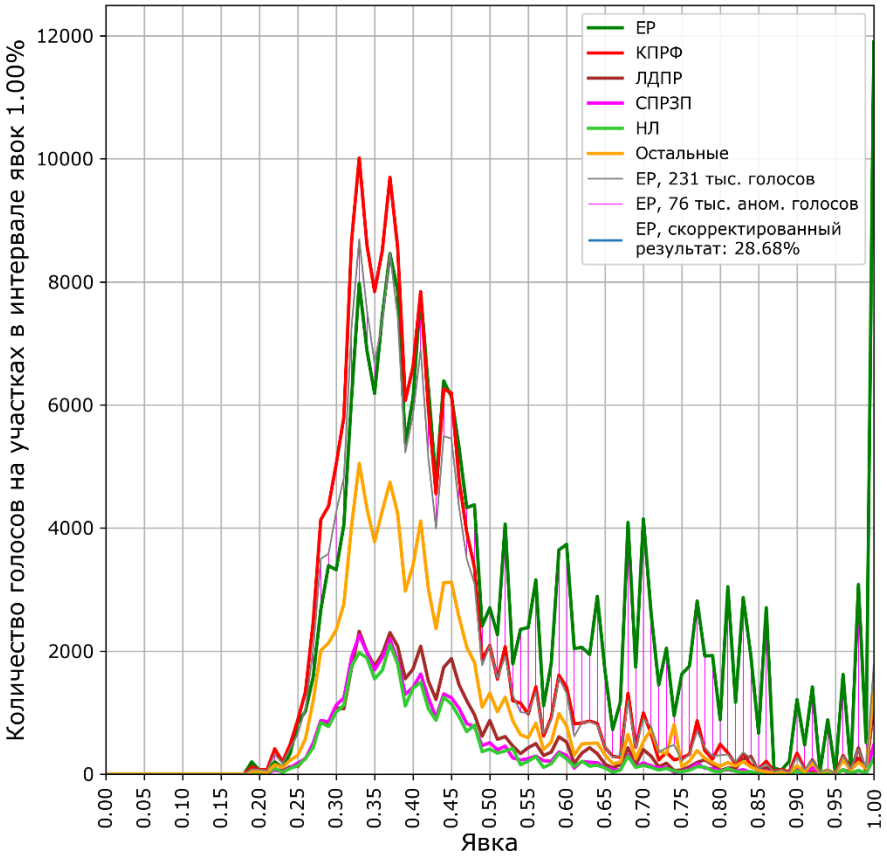


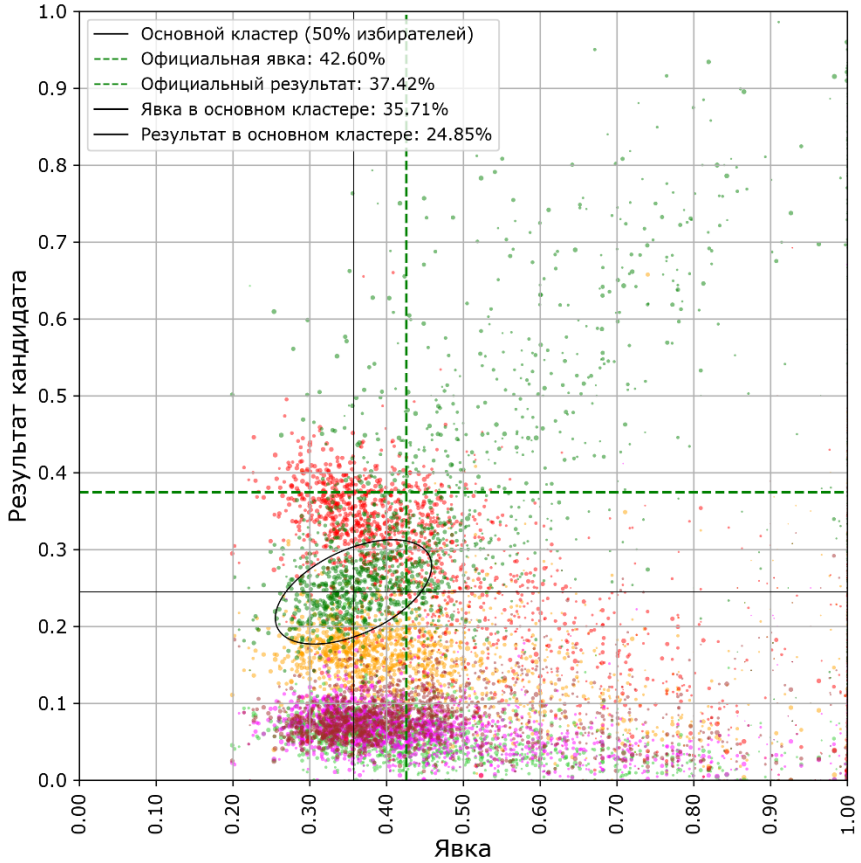
Республика Коми
Выборы в Государственную думу 2021 г.



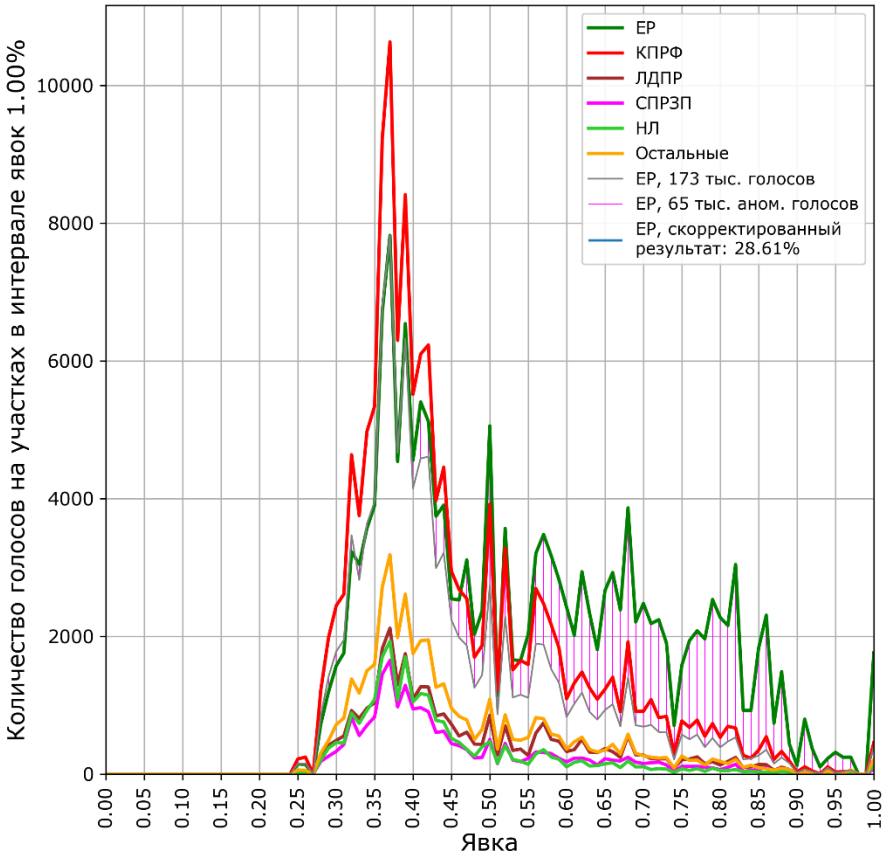


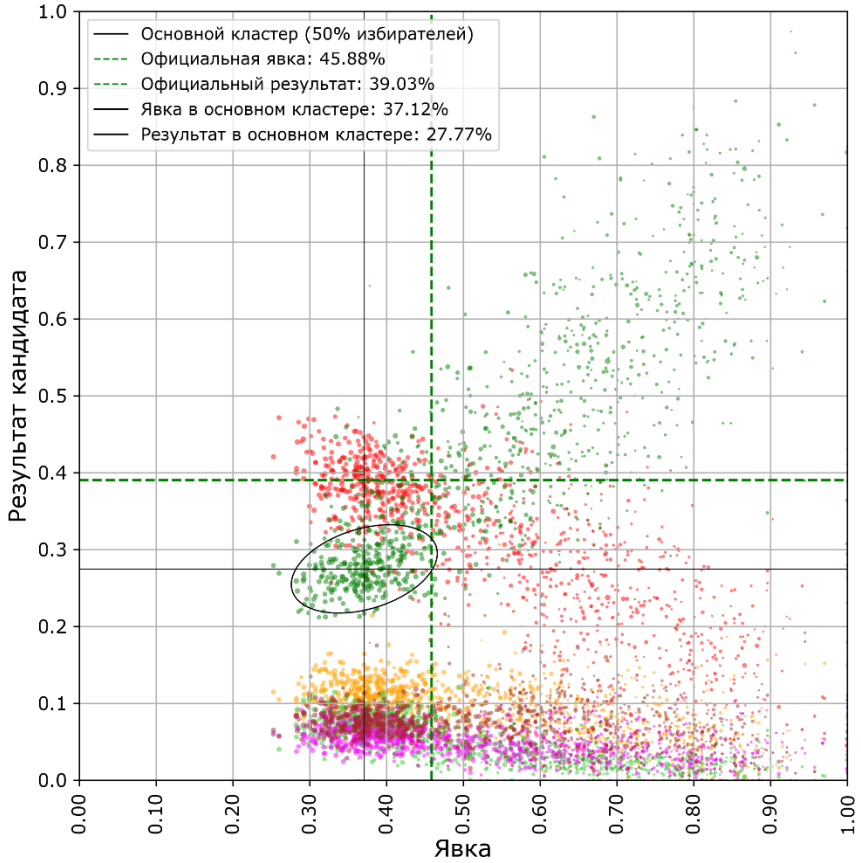
Приморский край
Выборы в Государственную думу 2021 г.



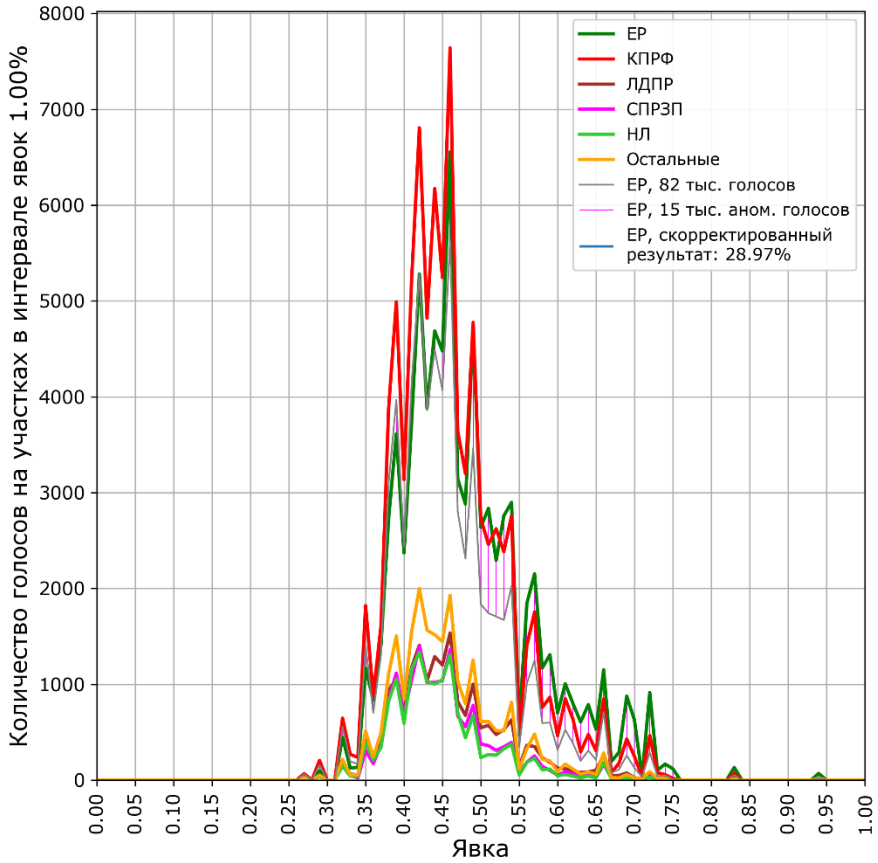


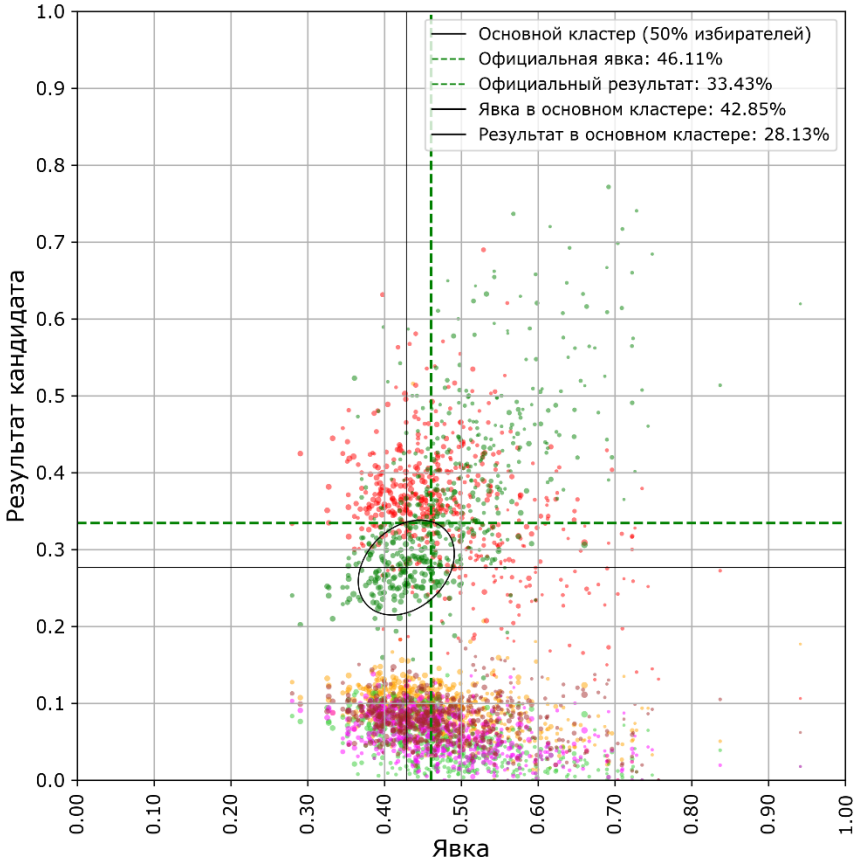
Ульяновская область
Выборы в Государственную думу 2021 г.



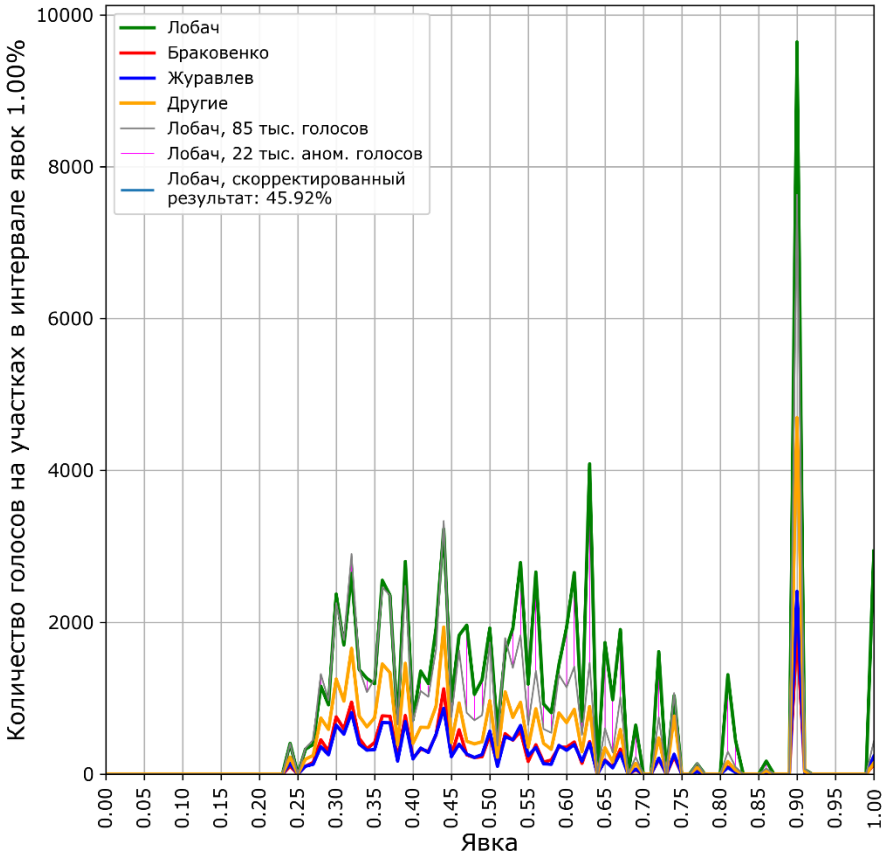


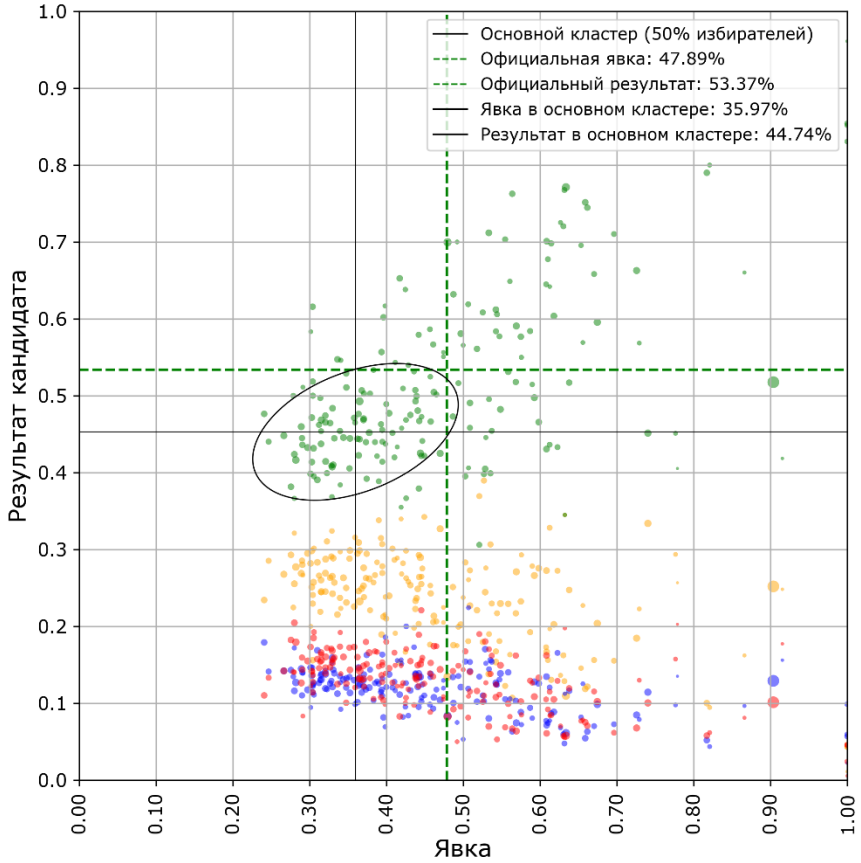
Республика Марий Эл
Выборы в Государственную думу 2021 г.



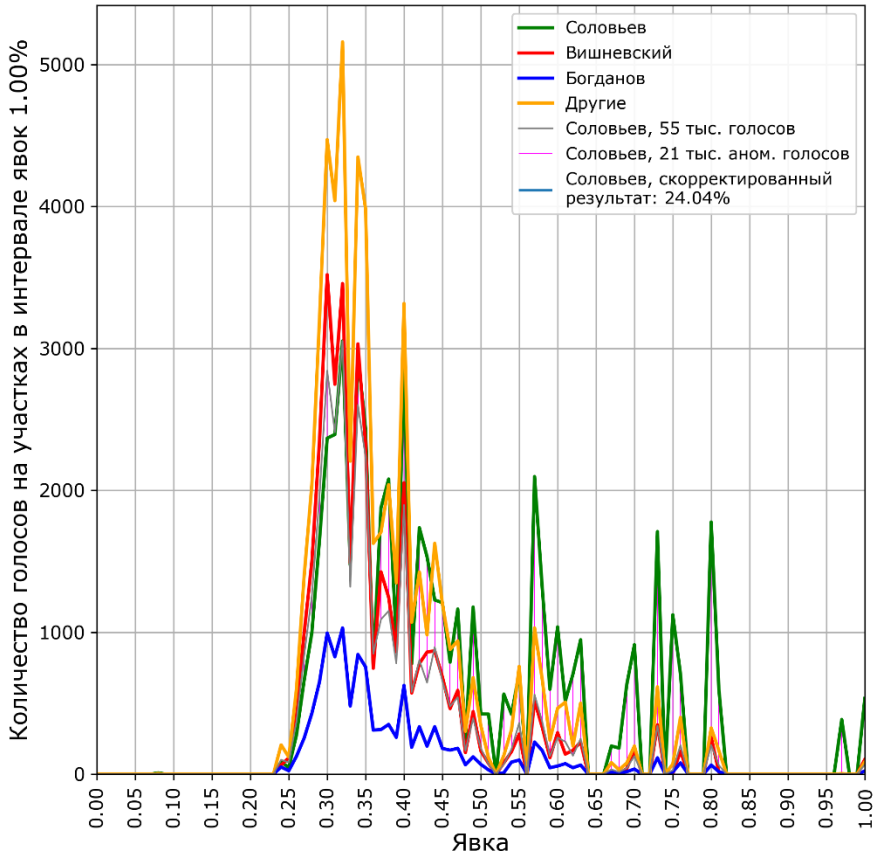


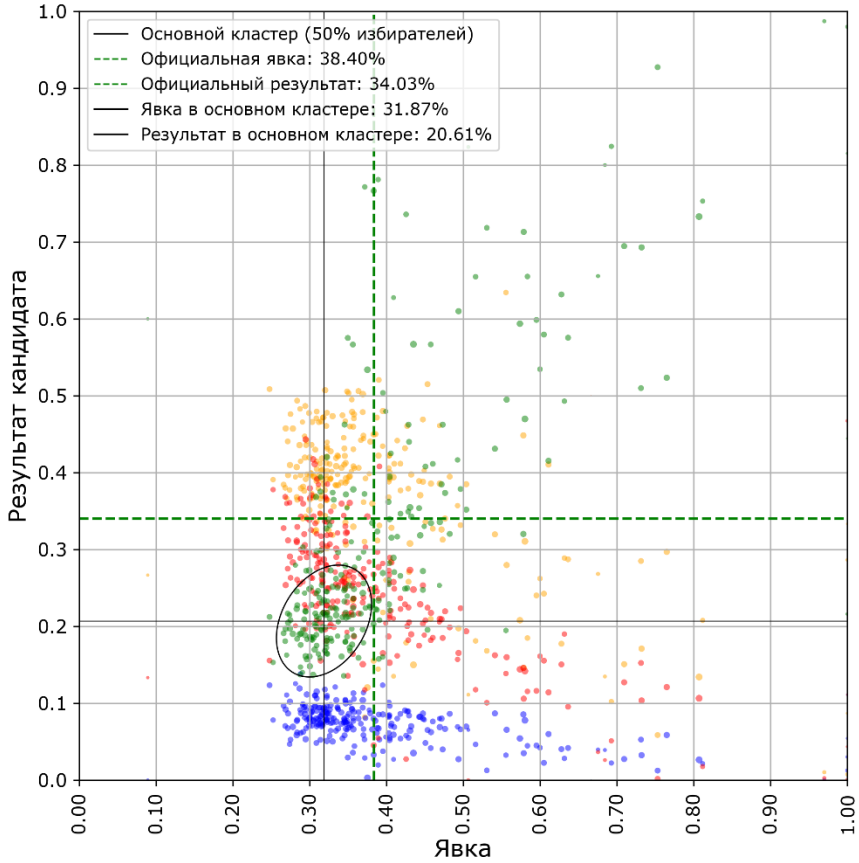
город Севастополь - ОИК №219



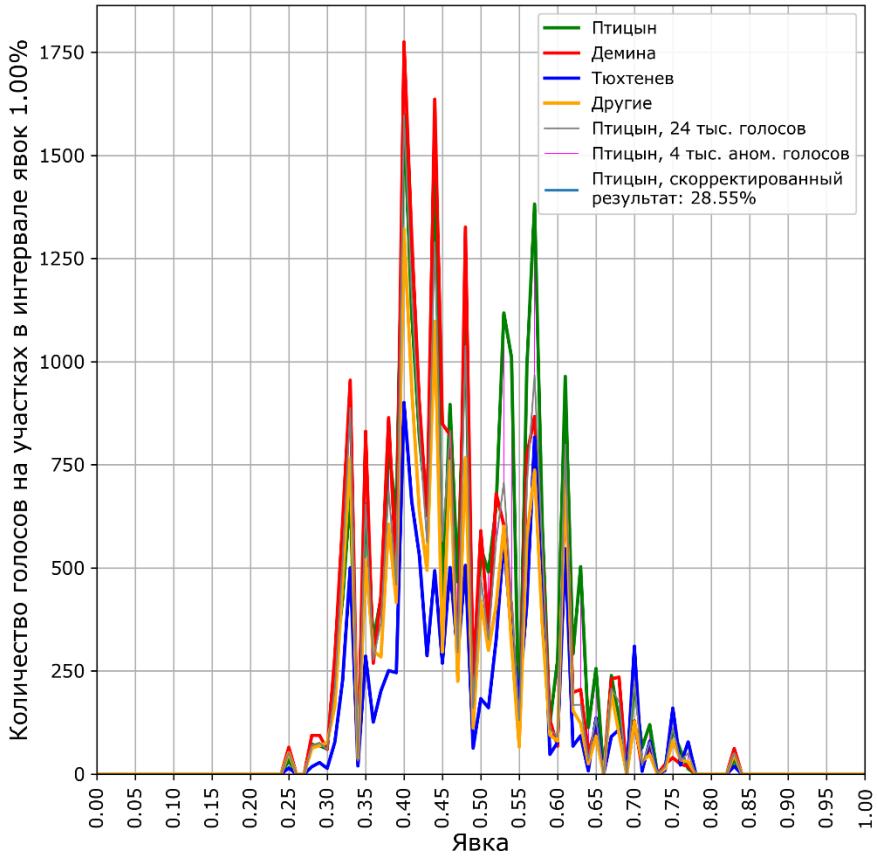


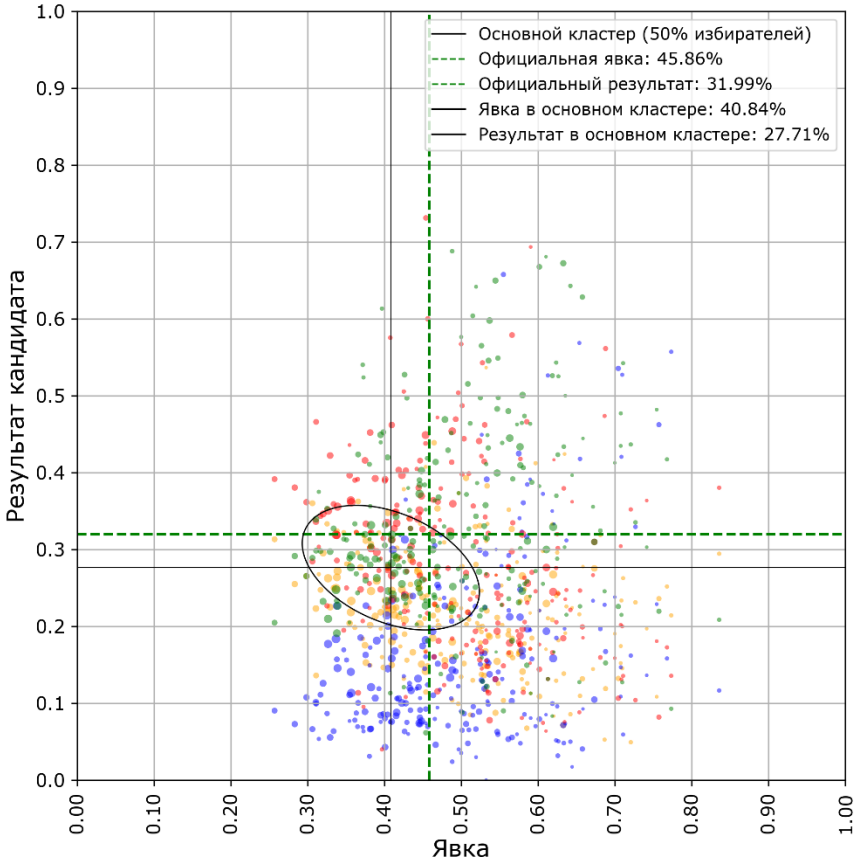
город Санкт-Петербург - ОИК №216



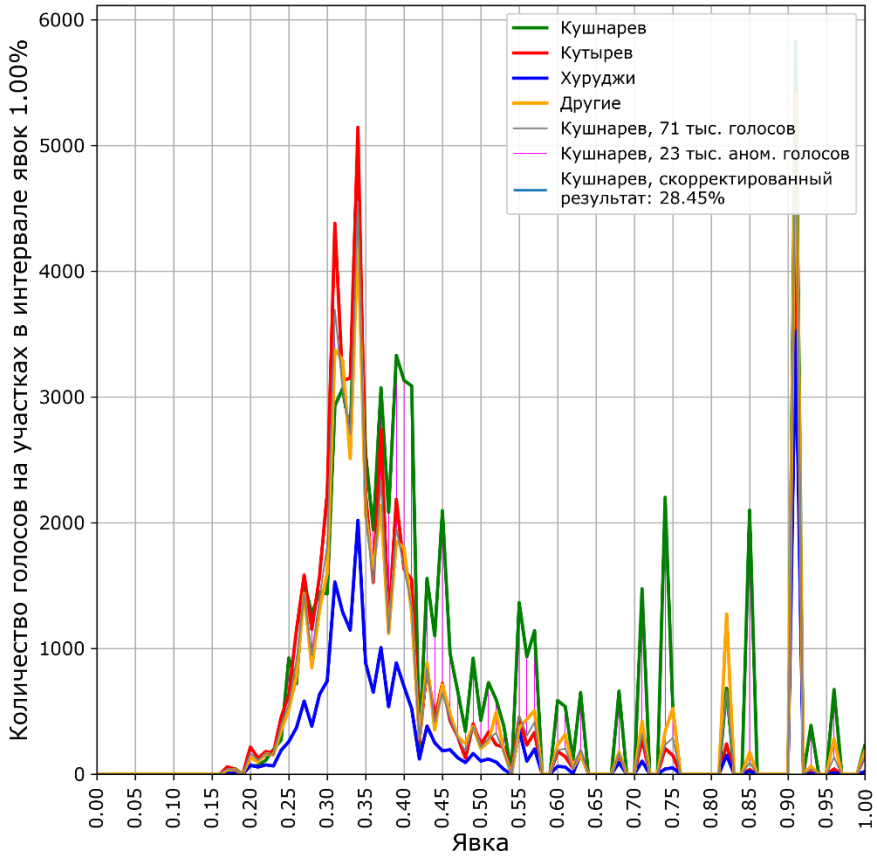


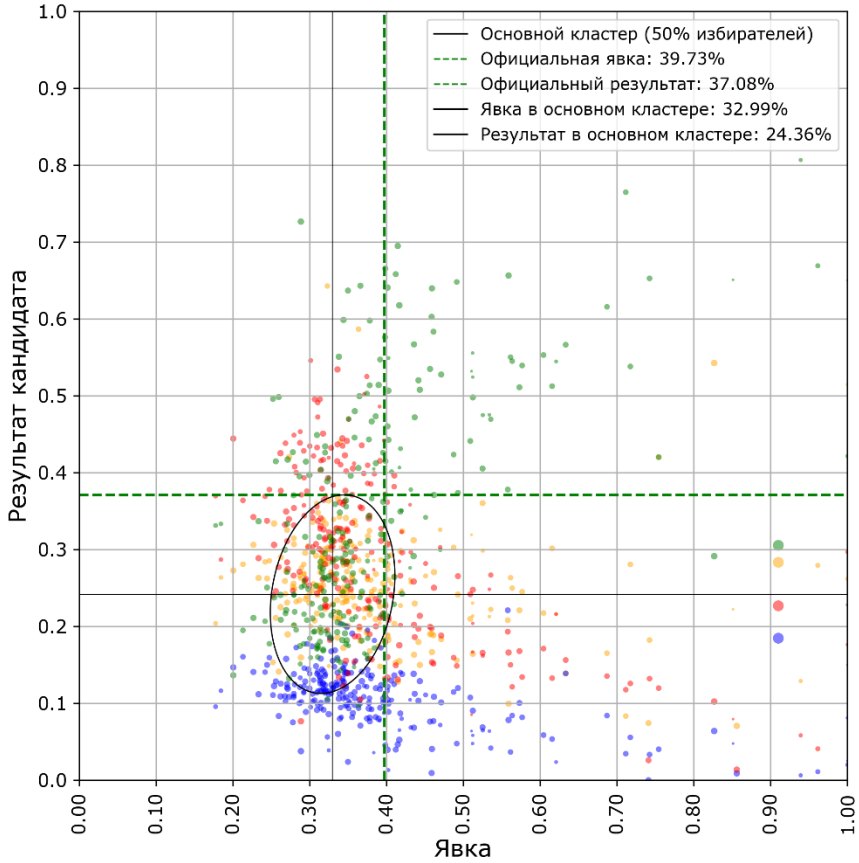
Республика Алтай - ОИК №2



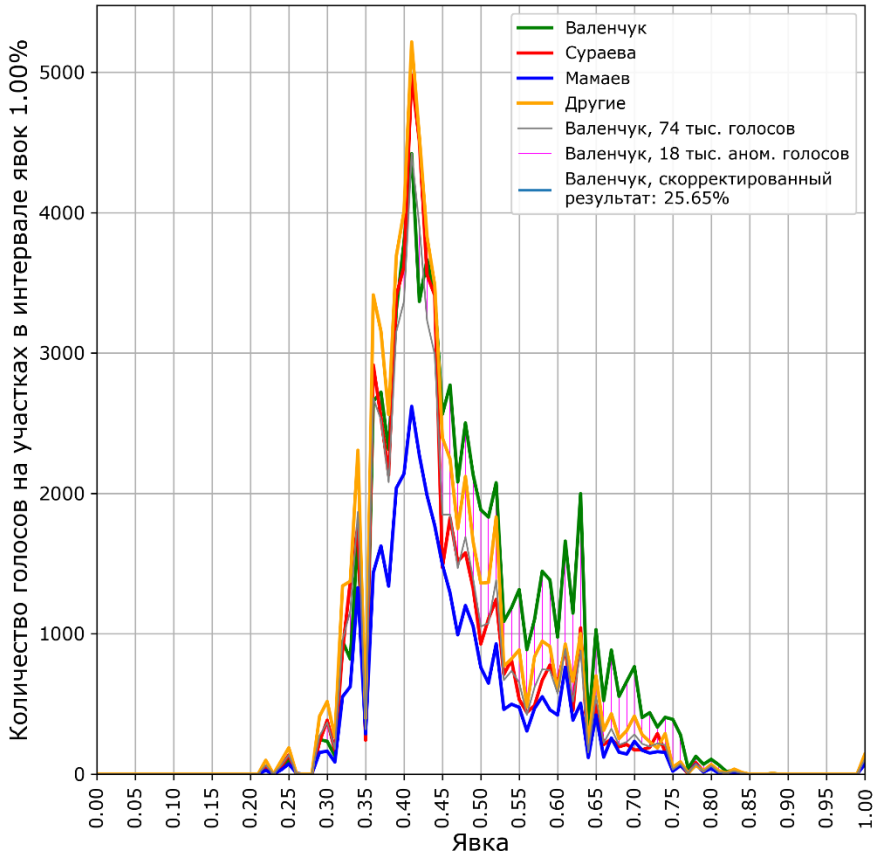


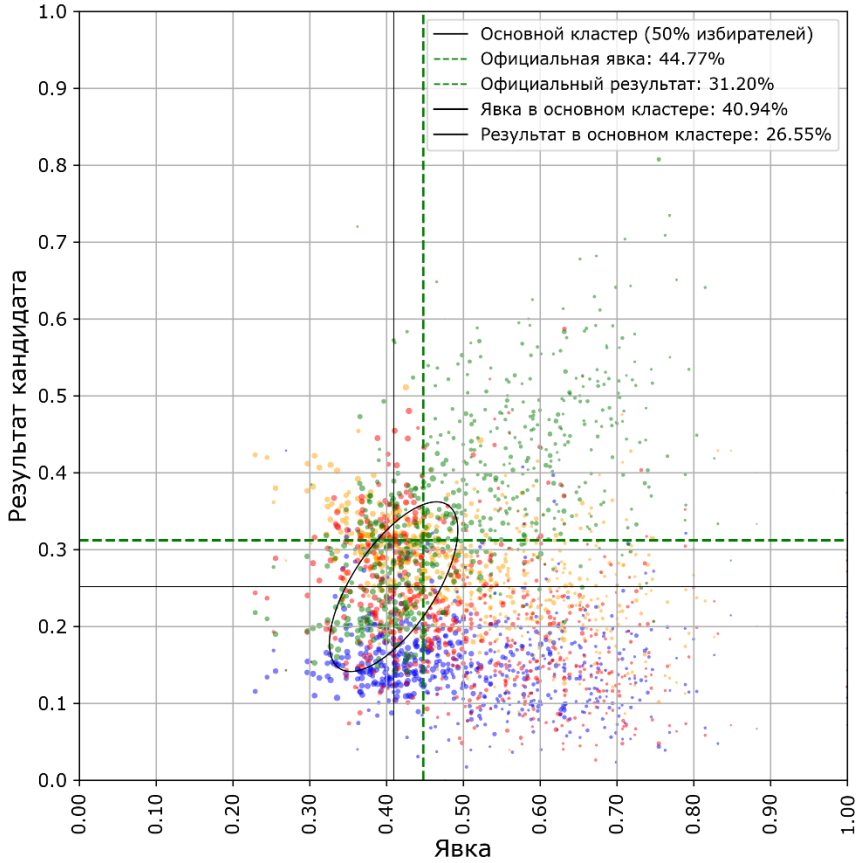
Ростовская область - ОИК №152



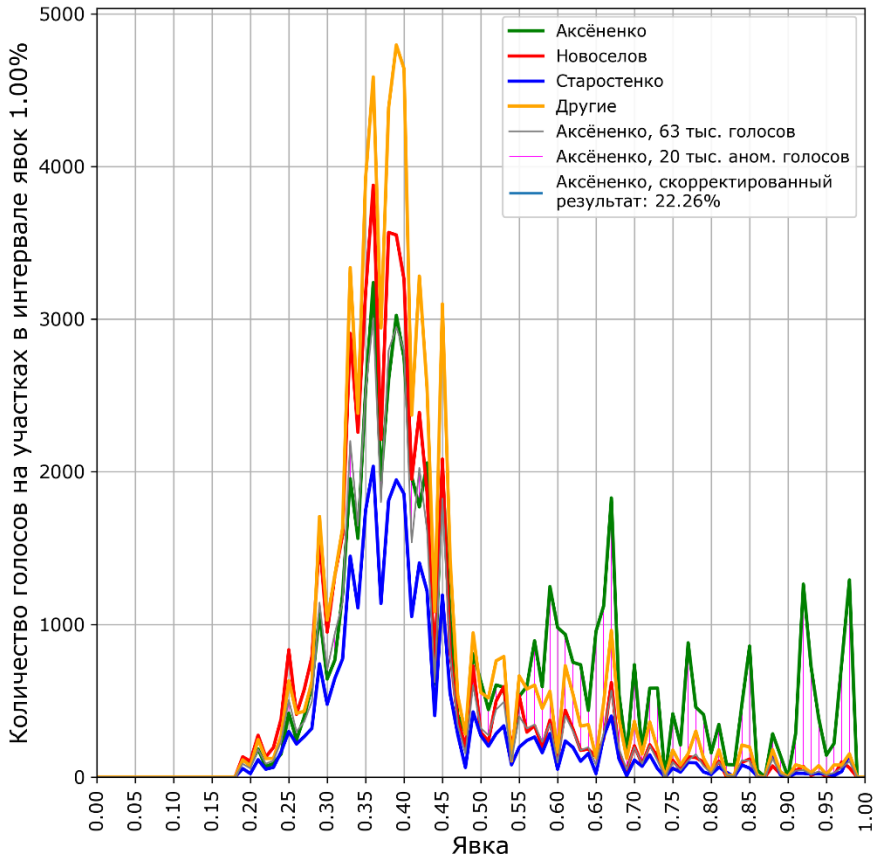


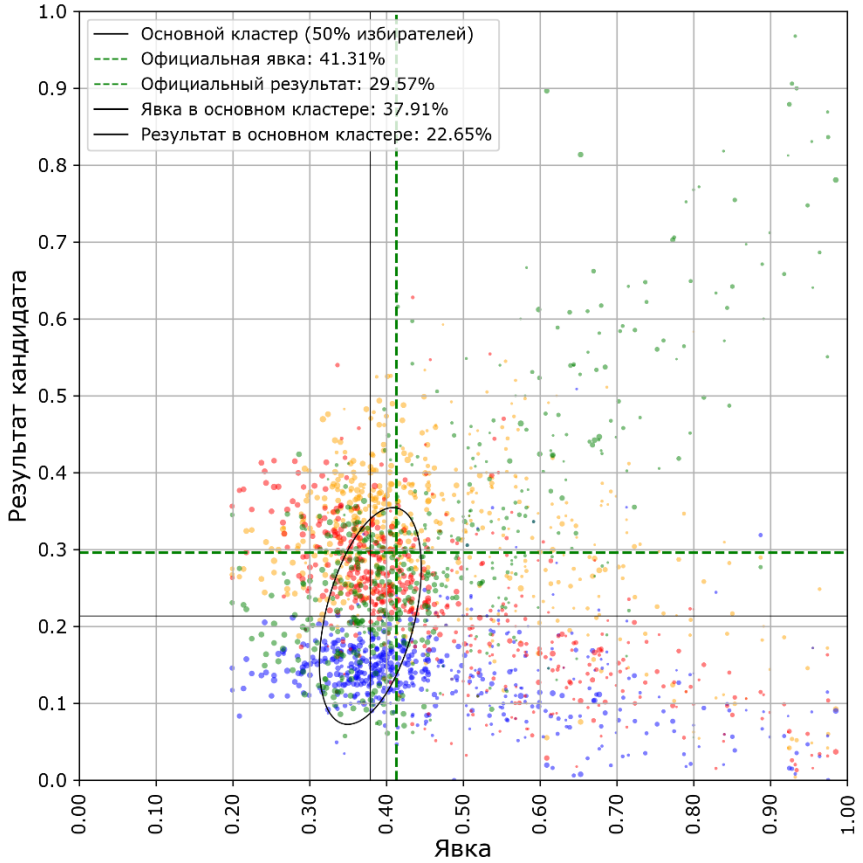
Кировская область - ОИК №106





Новосибирская область - ОИК №137





Статистические аномалии

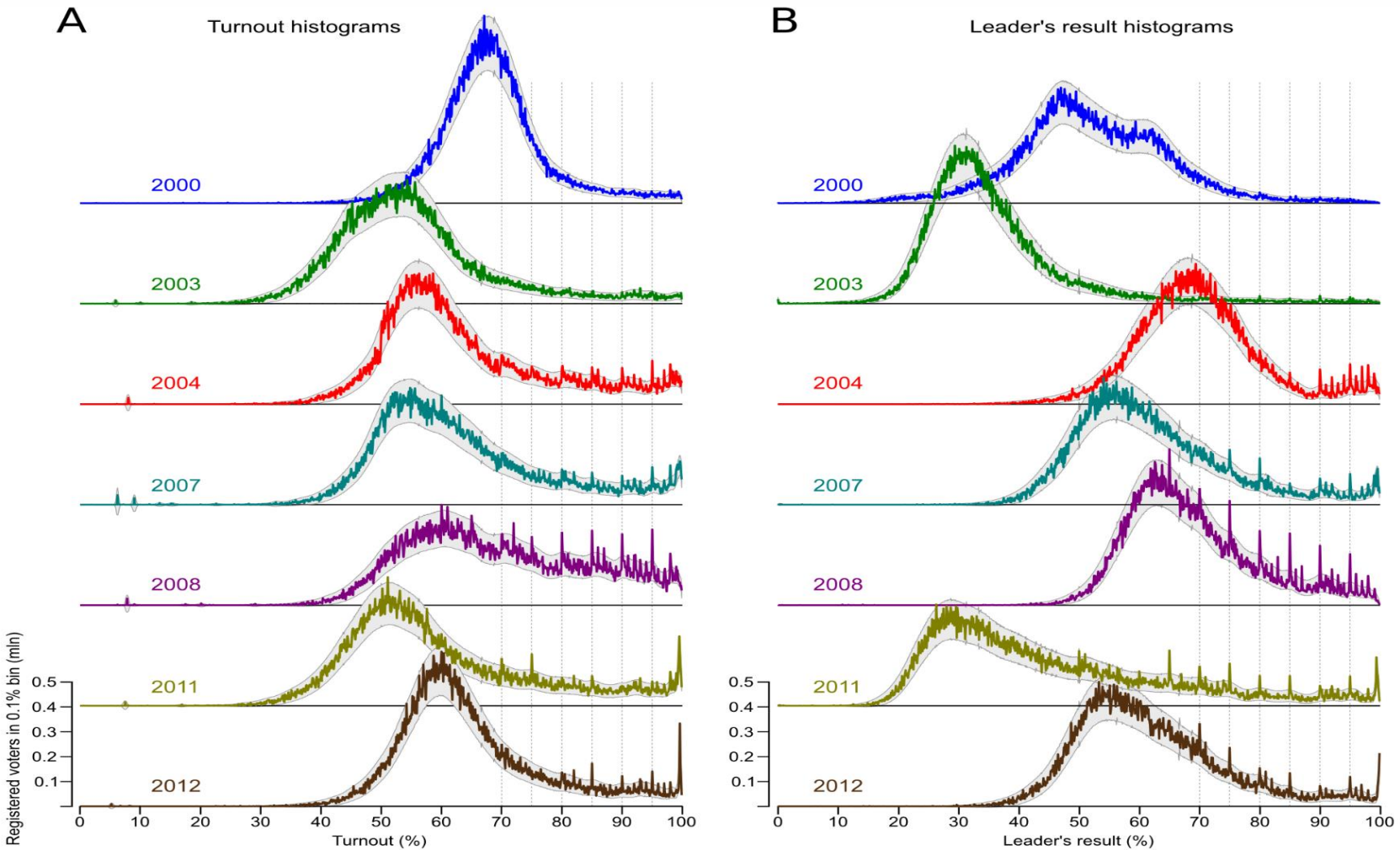
- Пики на целых процентах («пила Чурова»)
- Распределение последних цифр
- Числа под копирку
  - Вариант: «Калькулятор»

The Annals of Applied Statistics

Ann. Appl. Stat.
Volume 10, Number 1 (2016), 54-73.

Integer percentages as electoral falsification fingerprints

Dmitry Kobak, Sergey Shpilkin, and Maxim S. Pshenichnikov



Распределение последних цифр

- Распределение последней цифры для думских выборов 2011 г. http://trv-science.ru/2012/02/28/golaya-shkolnaya-matematika-2/

Числа под копирку

Северная Осетия, 2011 год, Левобережная ТИК Владикавказа

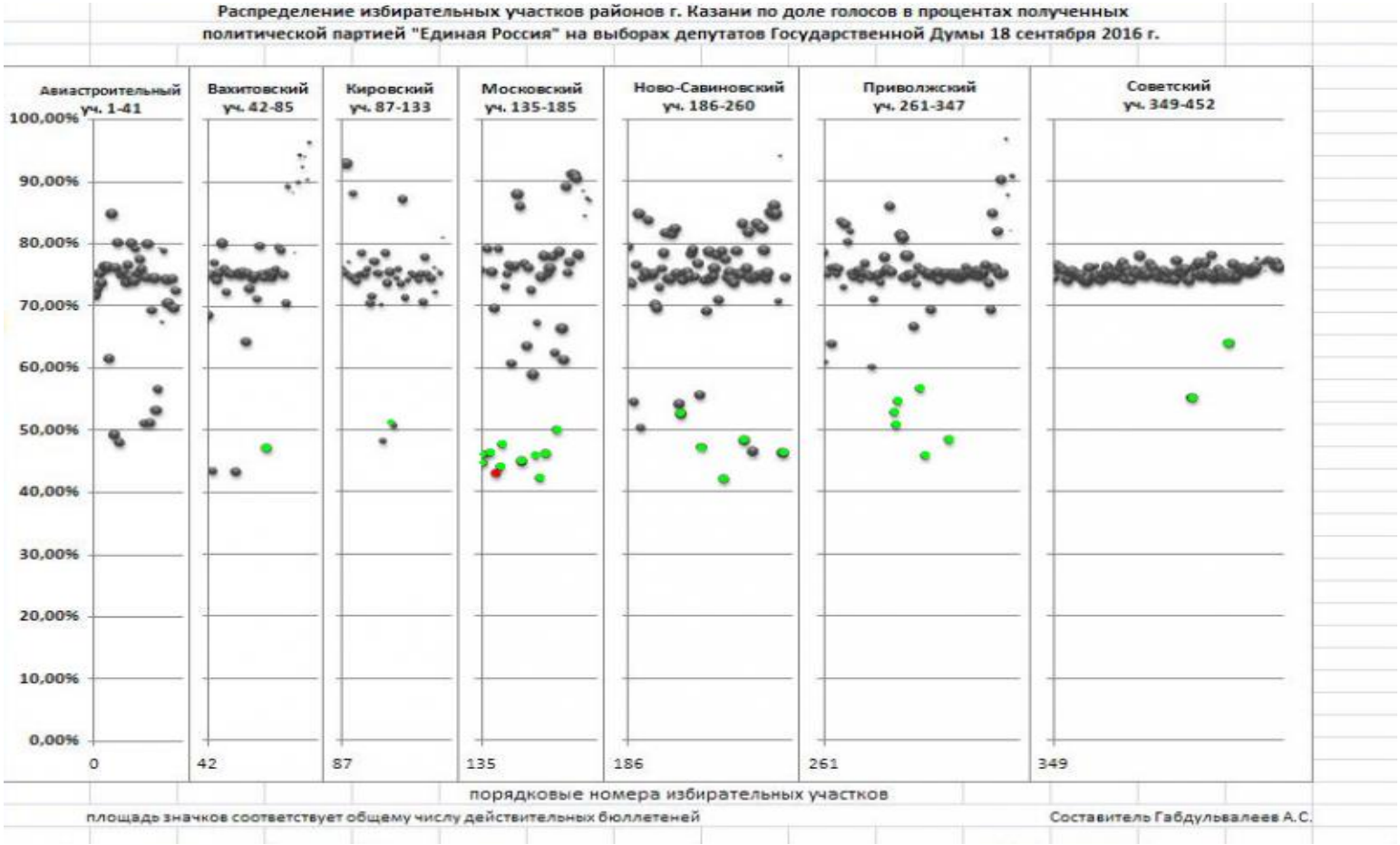
(https://kireev.livejournal.com/720076.html)

	Явка	СР	ЛДПР	ПР	КПРФ	Яблоко	ЕР	ПД
Сумма	91.71%	2.97%	1.09%	0.28%	24.96%	0.25%	68.88%	0.30%
УИК №299	89.46%	3.02%	1.15%	0.29%	20.11%	0.29%	73.28%	0.29%
УИК №300	92.32%	2.97%	1.07%	0.24%	19.36%	0.24%	74.11%	0.36%
УИК №301	98.69%	2.93%	1.06%	0.27%	19.79%	0.27%	74.00%	0.35%
УИК №302	92.00%	2.95%	1.09%	0.31%	20.19%	0.16%	73.60%	0.31%
УИК №303	96.17%	2.96%	1.06%	0.27%	19.60%	0.22%	74.03%	0.31%
УИК №304	88.74%	2.95%	1.04%	0.26%	19.72%	0.26%	73.76%	0.35%
УИК №305	91.57%	2.91%	1.07%	0.31%	21.17%	0.15%	71.78%	0.31%
УИК №306	92.81%	2.98%	1.19%	0.26%	19.15%	0.26%	74.30%	0.30%
УИК №307	91.92%	2.97%	0.98%	0.27%	19.80%	0.23%	73.83%	0.31%
УИК №308	88.80%	2.98%	1.11%	0.26%	18.82%	0.26%	74.70%	0.34%
УИК №309	91.28%	2.94%	1.06%	0.29%	20.26%	0.24%	73.42%	0.29%
УИК №310	83.94%	2.94%	1.10%	0.28%	19.69%	0.28%	73.87%	0.28%
УИК №311	91.96%	2.96%	1.10%	0.30%	19.73%	0.21%	73.98%	0.30%
УИК №312	92.27%	2.96%	1.10%	0.30%	74.12%	0.21%	19.79%	0.30%
УИК №313	89.70%	2.97%	1.08%	0.26%	19.79%	0.26%	74.18%	0.30%
УИК №314	88.31%	2.97%	1.08%	0.27%	19.68%	0.27%	73.84%	0.27%
УИК №315	93.19%	2.95%	1.14%	0.27%	19.72%	0.27%	73.91%	0.27%
УИК №316	86.67%	3.85%	0.00%	0.00%	19.23%	0.00%	76.92%	0.00%
УИК №317	86.39%	2.92%	1.17%	0.19%	17.70%	0.19%	75.49%	0.39%

Диаграмма Габдульвалеева

Г. Казань, 2016 год, явка избирателей

Зеленые точки – участки с наблюдателями



Литература

- Simpser, Alberto. 2013. Why governments and parties manipulate elections: Theory, practice, and implications. Cambridge University Press.
- Maxim Ananyev, Michael Poyker, Do dictators signal strength with electoral fraud?, European Journal of Political Economy, 2021, 102075
- Таблицы данных и диаграммы по регионам для федеральных выборов 1999-2018 гг: https://drive.google.com/open?id=0ByFMnUnpIlriRUl6ZWtSaER6eXM
- Kobak, D., Shpilkin, S. and Pshenichnikov, M. S., Integer percentages as electoral falsification fingerprints, Ann. Appl. Stat., v. 10, №1 (2016), 54-73. http://arxiv.org/abs/1410.6059
- Kobak, D., Shpilkin, S. and Pshenichnikov, M. S. (2016), Statistical fingerprints of electoral fraud?. Significance, 13: 20–23. doi:10.1111/j.1740-9713.2016.00936.x
- Kobak, D., Shpilkin, S. and Pshenichnikov, M. S., Statistical anomalies in 2011-2012 Russian elections revealed by 2D correlation analysis. https://arxiv.org/abs/1205.0741

Литература

- Borghesi C, Raynal J-C, Bouchaud J-P (2012) Election Turnout Statistics in Many Countries: Similarities, Differences, and a Diffusive Field Model for Decision-Making. PLoS ONE7(5): e36289. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0036289
- Kiesling J.B. Policing Armenian Elections: Turnout Distribution Analysis as a Tool for Quick Election Assessment and International Response
- Дискуссия «Двугорбая Россия» по итогам выборов 2016 г.; 20.10.2016 https://www.youtube.com/watch?v=Hs1fT4TJxQ0

Литература

- Исследование видеозаписей и анализ выборов на сайте Ассоциации наблюдателей Татарстана: http://tatobservers.ru/analiz-vyborov-2016-goda/
- Распределение последних цифр: http://trv-science.ru/2012/03/13/pro-arifmetiku-i-nemnozhko-pro-vybory/
- Анализ по недействительным бюллетеням: http://corbulon.livejournal.com/324525.html
- Анализ по малым кандидатам https://www.facebook.com/notes/boris-ovchinnikov/1015536824933230 4

Литература

- Серия статей о статистике выборов в Москве (в 5 частях): http://nonenoun.livejournal.com/4795.html
- Анализ дисперсии результатов по городам: http://trv-science.ru/2012/02/28/sto-vosemdesyat-chestnykh-gorodov/
- Выборы губернатора Приморского края 2018:
  - Расследование Ивана Шукшина по второму туру 16.09.2018: https://habr.com/ru/post/440092/
  - Аномалии результатов КОИБ-2010 16.12.2018: https://www.golosinfo.org/articles/143086
  - Подделка протоколов КОИБ-2010 16.12.2018: https://www.golosinfo.org/articles/143083

Литература

- Видеозаписи голосования 9 сентября 2018 г. во Владивостоке (участки 804, 820, 833, 837) – попробуйте найти момент подмены протокола https://archive.org/download/1809098331/1809098331_archive.torrent
- Доклад команды Р.Юнемана об электронном голосовании https://drive.google.com/file/d/1L9U2ssdjw_nRJMjBIzebhPDppfoWZgmJ/view